Rosa

Security and Compliance Guide

For customers and their security, privacy and third-party-risk reviewers.

Version	1.0
Issued	20 September 2026
Describes	The Rosa service as built and operated on the issue date, shared UK instance (eu-west-2)
Security owner	Jason Lee, founder
Approved by	Jason Lee
Next review	15 February 2027, or on any material control change
Contact	support@rosadebias.com

About this guide

This guide describes how Rosa is built and operated from a security, privacy and compliance perspective, so that you and your security or third-party-risk team can assess Rosa against the questions you would normally ask a data-processing vendor: how is my data isolated, how is it encrypted, who can access it, how long is it kept, how do you prove what you did to it, and what happens if something goes wrong.

How to read the status of any claim in this guide. Every control is marked one of three ways, and the distinction is deliberate:

Marker	Meaning
Evidenced	Operating, and backed by a dated test, configuration extract or verification we can show you
Asserted	Operating to the best of our knowledge, but not independently verified
Roadmap	Planned, not operating today

We do not describe a control as certified when it is not, and we do not describe a roadmap item in the present tense. Where a control is real but conditional, we state the condition rather than the headline.

Rosa is in a free validation phase, operated from the United Kingdom, and a UK limited company is being formed for the business. That stage shapes several answers below, most importantly around certification, contracting and formal service commitments.

What Rosa does with your data

Rosa removes encoded bias from tabular datasets. You submit a CSV and a small JSON configuration naming the protected attribute; Rosa returns a debiased dataset plus a report and an evidence record; a training job also creates a server-retained reusable transformation model, referenced by its training job id, which you apply to new data through an inference job. The model itself stays on the instance and is not a customer download. There are three operations: diagnose (measure bias), remove-bias training (learn a debiasing transform), and remove-bias inference (apply it to new data).

Rosa processes data on your instruction and does not sell, share or repurpose it. You decide what to send and why. The formal allocation of data-protection roles between us, per processing purpose, is being settled with legal advice and will be documented in the Data Processing Agreement at company formation. Until then this guide does not assert a role allocation, because a single blanket statement would be wrong: the position is not necessarily the same for your uploaded dataset as it is for account, access, billing, support and security-log data.

Rosa does not retain your dataset long term. Inputs, outputs and reports are removed on a short retention window described below. What is kept permanently is the Run Manifest: a record that a job ran, how it was configured, what it measured, and the integrity fingerprints of its inputs and outputs. It never contains the row-level dataset, the debiased output, or the model, but it does contain aggregate statistics that can include category labels from your data, set out under Data protection and privacy below.

During the validation phase we ask you to send synthetic or anonymised data, and to remove or pseudonymise direct identifiers (names, emails, account numbers) before upload rather than relying on configuration to exclude them. Note that pseudonymised data is generally still personal data under UK and EU law; the point at which Rosa is suitable for personal data is a contractual question we are taking legal advice on, and we will not onboard regulated production personal data without a Data Processing Agreement in place.

Architecture at a glance

Rosa runs as a small, isolated fleet on Amazon Web Services in the United Kingdom (London, eu-west-2):

- A hardened, always-on **front door** routes each request to the correct compute instance. It holds no customer credentials and stores no customer data. **It is a router, not the authentication boundary:** the compute instance authenticates the request and rejects an unrecognised key. A request that cannot be matched to a registered instance is routed to a default instance, where it is rejected.
- **Compute instances** do the processing. They start on demand and stop automatically when idle. Your uploaded file, the debiased output and the report are written to encrypted disk on the instance and remain there until the retention window elapses and the cleanup sweep runs: processing is short-lived, and storage is short but real.
- A dedicated customer can have their **own isolated cell**. The first dedicated cell is provisioned as part of an engagement rather than being available self-serve today. Running Rosa entirely inside **your own cloud account** (bring-your-own-cloud) is **planned for a future release**, not available today.
- An **evidence store** holds the immutable copy of each Run Manifest, on instances where the evidence vault has been provisioned.

Data protection and privacy

Retention windows. Your debiased output and your report are available to download for **7 days** after a job completes. Rosa also keeps an engine log on the same 7-day window and a trained model checkpoint for **365 days**;

both are internal files that Rosa manages on your behalf and are not downloadable - the checkpoint is kept so that inference against an earlier training job keeps working for a year, which Rosa resolves for you from the training job's ID. The raw input file is removed on the same 7-day window as the outputs.

How removal actually works, stated precisely. Access is refused the moment the window passes: a download of an expired artifact returns an explicit "expired" response whether or not the bytes are still present. The files are then deleted by a cleanup sweep that runs on the compute instance. **That sweep runs at most hourly and only while the instance is running**, and instances stop when idle, so in practice it runs about once per wake. The accurate statement is that access ends on time and the bytes are removed at the first sweep afterwards, not that files are deleted at exactly seven days.

What the sweep covers, and the one limit we disclose. The sweep covers the registered artifact files, the raw input file, the per-job working directory, and, on jobs which fail, are cancelled, or are interrupted, the engine diagnostic log that is preserved for support. Those diagnostic logs hold job identifiers, row counts and error detail rather than row-level customer data; they are customer-derived, so they sit on the same seven-day removal window as the outputs. The per-job working directory is removed once its files are gone, and a final sweep runs when an instance shuts itself down for idleness. **One honest limit follows from how the service is run rather than from a defect:** the sweep executes on the compute instance, so an instance that is never started again does not run it. **When an instance is retired, its encrypted storage volume is destroyed, which removes everything held on it** - a stronger removal than the sweep, not a weaker one. Access is refused on time in every case, and we will remove data on request at any point.

Evidence without keeping the data. Every job, whatever its outcome, produces a Run Manifest recording timestamps, the job configuration, and SHA-256 hashes of the exact inputs and outputs.

What the permanently retained Run Manifest holds, and where your data appears in it. Beyond identifiers, status, timestamps, versions, the served image digest and the hashes above, the manifest records: the declared columns and the dataset's size before and after encoding; **per-column statistics**, which for a categorical column are its distinct values and their counts, and for a numeric column its minimum, maximum and standard deviation; the **bias measurement** and its significance, and how it was measured; **per-group figures** for each value of the protected attribute, and, where a group was too small for the detection test to run, that group's value and row count; each feature's association with the protected attribute; on inference jobs, any **advisory warnings**, which can name a category or a numeric range from your data; and operational metrics about the run such as stage timings, peak memory and the hardware used. **So the manifest can contain category labels and column names from your data, and aggregate counts, but no row-level values.** This is why we ask you not to upload direct identifiers at all: listing a column as ignored keeps its values out of the manifest statistics, but the column is still uploaded and processed, so that setting is a transformation control and not a privacy boundary. *Evidenced: manifest schema and live manifests.*

What a hash does and does not establish. A SHA-256 hash lets you confirm that a file you hold is byte-for-byte the file Rosa recorded, and it ties that file to a specific job at a specific time. That is integrity and linkage, and it is genuinely useful in an audit. It does **not** prove that the debiasing method was correctly executed, that a stored record is immutable in storage (the write-once lock below is what addresses that, where it is provisioned), or that any deployment is compliant with a regulation. We are explicit about this because the difference is exactly what a competent reviewer will probe.

No customer dataset content in logs, and none in any language model. Operational logs record job identifiers, row counts, event outcomes and error codes. Access logs additionally record a source address and a non-reversible key fingerprint, which are personal data and sit under the retention and residency controls above. Customer CSV content is not written to logs, and customer data is never passed to a large language model: Rosa's debiasing core contains no language model and makes no external model calls. Support diagnostic bundles are assembled from the same operator logs and carry no row-level customer data.

GDPR. Rosa is engineered to GDPR principles now: minimised and short-lived processing, UK data residency, and a documented control set. The formal programme, being a Data Processing Agreement, a published privacy notice, an Article 30 record of processing, and a documented data-subject-rights procedure, is executed at company formation, planned for late 2026. For production personal data, the dedicated-instance or bring-your-own-cloud path is the appropriate route.

Data residency

All processing and all storage for the shared instance are in **AWS eu-west-2 (London, United Kingdom)**. eu-west-2 is in the United Kingdom, which since Brexit is a third country for EEA transfer purposes. If you are an EEA-established customer, transfers to the UK currently rely on the European Commission's UK adequacy decision. We will name the relevant transfer basis and any subprocessors in the Data Processing Agreement. Rosa does not copy customer data to another region.

Encryption

In transit (public). All public API traffic is served over TLS 1.2 or higher, in practice TLS 1.3 with AES-256-GCM, with HSTS and an automatic HTTP-to-HTTPS redirect. Certificates renew automatically. *Evidenced.*

In transit (internal). Every internal service-to-service hop is mutually authenticated and encrypted with TLS across the fleet. Each service presents its own certificate and verifies its peer against a certificate authority unique to that deployment cell. The certificate authority's signing key is held in AWS Key Management Service, backed by validated hardware security modules, and is not exportable. Certificates rotate automatically with expiry monitored in advance. This runs on top of a private, security-group-restricted network. *Evidenced.*

At rest. Every storage volume across the fleet is encrypted with AES-256 using AWS-managed keys, verified per volume. The evidence archive in object storage is server-side encrypted and reachable only over TLS from within the private network. *Evidenced.* **Customer-managed keys (bring-your-own-key) are not available today.** *Roadmap, dedicated tier.*

Browser surface. The customer portal serves a pinned, self-hosted JavaScript runtime under a restrictive Content-Security-Policy that permits network connections only to Rosa's own API, so a script running in the portal origin cannot transmit your API key elsewhere. *Evidenced.*

Two disclosed limitations of that policy, and what we do about each. We would rather state these than let a reviewer find them.

First, the policy permits `unsafe-eval`, and we cannot currently remove it. The portal's UI framework evaluates its display expressions at runtime, which is what that permission allows. The framework does publish a build that does not need it, but that build forbids constructs the portal uses throughout, so adopting it is a rewrite of the interface rather than a configuration change. **Two independent reviews have raised it, and we accept it rather than claiming otherwise.**

The acceptance is monitored rather than permanent, which is the part that matters for assurance. An accepted risk with nothing watching it degrades quietly as code is added; this one fails the build if a new dynamic-HTML injection point appears anywhere in the portal. **The current inventory, measured rather than asserted:** the portal application itself contains **no** dynamic-HTML injection points fed by data; its seven templated bindings render hardcoded labels, numeric values checked and clamped before display, and a fixed glossary. **The downloadable Data Checker contains nine.** Five of them render values derived from your file - column names, the verdict text and every finding - and each such value passes through one shared HTML-escaping function before it is inserted. Three render only numbers (the row and column counts, the encoded-width figures and the sampling note), which are coerced to a number before display, a stronger guarantee than escaping. The ninth renders fixed text only. **So no value derived from your file reaches an injection point unescaped** - though a reviewer counting calls to the

escaping function will find it at five of the nine sites, not all nine, because the other four never handle a value that could carry markup. The Data Checker also never transmits your data: it runs entirely in your browser.

We state the nine explicitly because the count is easy to get wrong: the portal application has none, and the download has nine, so a figure quoted for "the portal" alone understates it. A control described from the wrong count is worse than one described honestly. *Asserted, with the guard Evidenced.*

Second, the policy is delivered in the page rather than as a response header, which also means the portal cannot currently send HTTP Strict Transport Security, `X-Frame-Options`, `X-Content-Type-Options` or `Referrer-Policy`. **This is a property of the static hosting the portal runs on, which provides no mechanism to set response headers at all - it is not an omission in our configuration, and no change to our code can add them.** The remedy is a hosting change, which is on our list and is in scope for the pre-listing security review. **In the meantime the connection is HTTPS-only and the API-egress restriction above is unaffected**, because that part of the policy works identically whether it is delivered in the page or in a header. *Roadmap.*

Tenant isolation

Workspace-scoped access. Every data access is scoped to your workspace identifier, enforced at the storage layer and covered by negative-isolation tests that assert one workspace cannot read another's jobs, artifacts or manifests. On 6 September 2026 the same boundary was also tested adversarially at the API layer against the running shared service, using two real workspace keys on one instance: reading another workspace's job record, manifest and artifact list, downloading its files by exact filename, cancelling its job (a write), and submitting under an idempotency key it had already used were all refused, nine checks of nine, while the same operations on the requesting workspace's own jobs succeeded. That test exercises the API boundary, not the storage layer, and it is an internal test rather than an independent one. *Evidenced.*

A routing fault is designed to be a denial, not a disclosure. Authentication happens at the compute instance, not the front door, so an instance that receives a request bearing a key it does not hold rejects it rather than serving it - and every cross-workspace refusal answers 404 rather than 403, so the existence of another workspace's resource is not confirmed to the caller. This is the designed control against the broken-object-level-authorisation class of failure. We describe it as designed and tested rather than as a guarantee that no path exists, because no vendor can honestly assert the absence of every path, and an independent penetration test has not yet been performed. *Evidenced as designed and tested; independent verification is Roadmap.*

Isolation you can dial up. The shared environment is logically isolated. A dedicated cell provides a separate front door, hostname, network boundary, backup set and blast radius. Bring-your-own-cloud, which would run Rosa inside your own AWS account, is **planned for a future release**. *Dedicated cell: offered, provisioned per engagement. Bring-your-own-cloud: planned, not available today.*

Identity and access management

Customer access is by API key. **Plaintext customer keys are not persisted server-side and are not written to logs**; what is stored is an argon2id verifier hash, and the routing layer keys on a non-reversible fingerprint. A key exists in plaintext where it necessarily must: in your possession, in the request over TLS, and transiently in memory during verification. We state it this way rather than "no key material exists anywhere", which would not be true of any authentication system. Keys can be rotated and revoked at any time, and a revoked key stops working immediately across the service. *Evidenced.*

Provisioning is operator-controlled. There is no anonymous or self-serve path to mint a credential; workspaces are created only through an authenticated administrative path. *Evidenced.*

Administrative access is protected by a hardened, rotatable admin secret held in root-only storage on each instance, separate from the code and never in version control. Cloud administrative permissions are least-privilege:

the day-to-day automation credential is scoped to what it needs, and the ability to create or destroy infrastructure is held separately behind multi-factor authentication with no long-lived access keys. *Evidenced.*

Roadmap for enterprise customers: single sign-on federation (SAML / OIDC), enforced key rotation and expiry, mutual-TLS client certificates, and scoped read-only keys.

Availability, backups and resilience

Backups with a tested restore. All fleet volumes are snapshotted on a managed schedule, 14 daily and 12 monthly. The restore has been rehearsed and dated: on **14 July 2026** a snapshot was restored to a scratch volume, both databases opened, and `integrity_check` returned `ok` with the workspace and job tables intact and a Run Manifest parsed successfully. The rehearsal validated recoverability; **it did not measure a recovery time**, and a formal recovery-time objective has not yet been set. *Evidenced for recoverability; recovery-time objective is Roadmap.*

Why backups are kept longer than data. Snapshots of each instance's storage are retained on a rolling schedule of 14 daily and 12 monthly. The twelve-month tier is deliberate: trained models are available to you for 365 days, and a backup has to outlive the longest commitment it protects, so that a hardware failure does not cost you a model you are still entitled to use. A snapshot is an image of the whole volume, so short-lived data that happened to be present when it was taken is carried in that image until the snapshot rotates out of the set. Those copies are not reachable through the service, are never searched or analysed, and exist solely to rebuild an instance. If you ask us to erase your data we remove it from the live service immediately; any copy inside a backup expires with that backup's cycle and is not restored back into use. *Evidenced for the schedule and its purpose.*

Off-site immutable evidence, and its condition. Each Run Manifest is copied off the machine to separate object storage under a seven-year write-once (Object Lock, COMPLIANCE mode) retention, which refuses deletion including by an administrator, verified against the live bucket. **This applies to instances where the evidence vault has been provisioned.** The shared instance has one, so it covers all current customers. A newly stood-up dedicated or bring-your-own-cloud instance is durable from the outset (snapshotted volume plus an hourly state mirror) but is not write-once-immutable until its vault is provisioned. Ask us to confirm the status for your instance. *Evidenced where provisioned; not claimed otherwise.*

Contingency for continuity. The serving images and the key and job state are mirrored to durable storage on an hourly cycle, bounding potential state loss to about one hour, so a machine can be rebuilt from a known-good source. *Evidenced.*

Monitoring and alerting. Job failures, timeouts, application errors and low disk raise automated alerts, and the front door is monitored by an independent external prober. *Evidenced.*

Abuse protection. Requests are rate-limited per key and per source address, and oversized uploads are rejected before any data is written to disk. A stopped compute instance is started only for an authenticated request with a key the front door recognises, and at most a set number of times an hour from any one source address (currently 20 on the shared instance); past that, the request is refused and nothing is started, which bounds the cost an abusive or misbehaving client can cause. *Evidenced.*

Honest scope. Rosa runs in a single region without multi-region failover. There is currently **no contractual service-level agreement, no committed incident-response time, and no formal change-management commitment to customers**; these are defined per engagement and formalised at company formation. For a dedicated engagement we agree recovery objectives explicitly.

Secure development and change management

- **All changes ship through peer-reviewed pull requests and continuous integration.** Nothing reaches the main line directly. *Evidenced.*

- **Dependencies are scanned automatically** on every change, with builds blocked on fixable vulnerabilities and a 72-hour internal target for critical fixes. This is an internal target, not a contractual commitment. *Evidenced.*
- **Secrets are scanned automatically** across the codebase and its full history, backed by a documented secrets inventory and rotation policy. *Evidenced.*
- **Releases are versioned and provenance-stamped.** Each instance records the exact image it serves in every Run Manifest it produces. Customer-serving instances are pinned to an exact version and move only when an operator updates them deliberately; the shared fleet's own default instance, which never carries a customer workspace, tracks the current release instead. *Evidenced.*

Logging, audit and monitoring

Rosa keeps two structured log streams, an application event log and a per-request access log, centralised with 365-day retention. Each access record carries a non-reversible key fingerprint, never the key, plus the endpoint, method, UTC timestamp, source address, outcome and a correlation id. Logs are written append-only and the serving instance holds append-only permission to the central store, so an instance cannot retrospectively rewrite its own audit trail. *Evidenced on mechanism.* **One limit, stated plainly: the buffer that ships records to the central store is bounded.** Under sustained pressure it discards the oldest queued records rather than growing without limit, and those records are lost. This is a deliberate trade - a bounded, discoverable failure in preference to an unbounded buffer that could exhaust the instance and take the service down with it. Discards are counted in the process, but that counter is not yet exposed on any endpoint, so a gap is currently identifiable after the fact rather than alerted on. We expect this only under abnormal load, and a gap in a stream should therefore not be read as evidence of an outage without checking. Customer log export into your own SIEM is *Roadmap, dedicated tier.*

The evidence layer

Rosa is an evidence-producing service as well as a processing one. A completed job emits a Run Manifest, and diagnose and training jobs additionally produce a PDF report. Not every job produces both. The full matrix:

Outcome	PDF report	Debiased CSV	Run Manifest
Diagnose, complete (including a diagnose that found no recoverable bias)	Yes	No	Yes
Remove-bias training, complete	Yes	Yes	Yes
Remove-bias inference, complete	No	Yes	Yes
Remove-bias training, declined (no recoverable bias)	Yes	No	Yes
Diagnose or remove-bias training, detection not supported	Yes	No	Yes
Failed, cancelled or interrupted	Depends where in the run it stopped	No completed output	Yes

A declined job is worth calling out: it produces a complete evidence report explaining what was tested and why no recoverable bias was found, not merely a record that it stopped.

"Detection not supported" is a different outcome and must not be read as a clean result. When any group of the protected attribute has fewer than nine rows in the training sample, the detection test is not run, because a group that small cannot be scored on a held-out split. The job ends with that status, for diagnose and training alike, with an evidence report and a Run Manifest naming the group and its row count, and no debiased output. **It carries no verdict and says nothing about whether the data is biased.** A record showing it is a record that the test could not be performed, and a reviewer should treat it as such. *Evidenced: service behaviour and manifest schema.*

These outputs **contribute to** data-governance evidence that emerging AI regulation asks operators to produce. They do not on their own satisfy any of these frameworks:

- **EU AI Act Article 10** (data governance for high-risk systems). Rosa supplies the bias-examination and per-run traceability part of the record. Article 10 also covers design choices, data provenance and collection, preparation, stated assumptions, suitability and representativeness, which are yours to evidence. Note also that the obligations differ by scope: Annex III high-risk systems and Annex I product-embedded systems are on different tracks and different dates, so check which applies to you rather than reading a single date.
- **ISO/IEC 42001** (AI management systems). Rosa's outputs map to data-quality assessment, risk-register entries, traceability and release-gate evidence within your management system.
- **NIST AI Risk Management Framework.** The same outputs align to the Govern, Map, Measure and Manage functions.

Rosa is best understood as a data-centric assurance layer that helps you evidence your own controls with repeatable, per-run records.

What a bias result establishes

Rosa's bias and residual figures are produced by one instrument: a capacity-limited discriminator scored as a threshold accuracy on a held-out split, calibrated against a permutation null. Reported honestly, that instrument establishes that **i**t did not recover the protected attribute above chance at the recorded sample size, repeat count and permutation resolution. It does not establish that no estimator could. Independent reviewers running their own recovery models on our published benchmarks have found signal above chance in outputs this test reported as clean, and we treat that as the bound of the measurement rather than a contradiction of it. *Evidenced as a bounded measurement; not a guarantee of statistical independence.*

Two measurement improvements are specified for the next engine generation: a rank-based detection statistic reported alongside the threshold accuracy, and a prespecified external recovery adversary, independent of the Rosa codebase and of a different model family, whose result is published beside the internal figure. Neither operates today. *Roadmap.*

Until they do, the reduction between the before and after measurements is the part of the report that is directly comparable, because both sides are measured by the identical instrument. If the strength of a fairness claim matters to a decision you are accountable for, run your own recovery model on the debiased output and let that number inform it.

Bias removal and utility preservation are separate findings

A clean residual is not a statement about whether the debiased data remained useful. They are independent axes, and we have measured them diverging: in a controlled campaign at small sample sizes, the manifest residual read *not detected* on every run **including the runs whose downstream utility ordering failed**, and post-debias recoverability was flat across passes and failures alike. Nothing in the report or the manifest distinguishes those two outcomes.

The practical consequence for a reviewer: **the evidence Rosa produces establishes what was done to the data and what its own test could recover afterwards. It does not certify that a model built on the output performs acceptably.** That has to be established on your side, on a held-out split, against your own success criteria - which is why we recommend pre-registering that acceptance test before the run rather than reading it off the manifest afterwards, and why the published no-trade-off benchmark is stated with an explicit minimum sample size rather than as a general property. *Evidenced as a known and disclosed limit of the evidence layer.*

Compliance posture

- **SOC 2.** Rosa is designed toward the Security and Confidentiality criteria and operates the corresponding controls. **No SOC 2 audit has been performed.** A readiness assessment and, subsequently, an independent penetration test are planned ahead of the first regulated engagement. *Roadmap.*
- **Independent penetration test.** Not yet performed. *Roadmap.*
- **GDPR.** Principles applied now; formal programme executed at company formation. *Partly Evidenced, partly Roadmap, as set out above.*
- **ISO/IEC 42001 and NIST AI RMF.** Rosa's outputs are mapped to both and are usable within your own programme today. Rosa holds no certification against either. *Asserted mapping; certification is Roadmap.*
- **EU AI Act.** Rosa produces evidence that contributes to Article 10 obligations. High-risk obligations under the current timetable apply from 2 December 2027, with the Annex I track on a separate schedule. Rosa does not make you compliant. *Asserted.*
- **ETSI EN 304 223** (AI cybersecurity baseline). Rosa is designed to support organisations implementing aligned controls at the data-governance layer. No certification is claimed. *Asserted.*

Bank and regulated-sector readiness. Two prerequisites remain open and we will not imply otherwise: an independent assurance report and penetration test, and the corporate and contractual layer (the executed Data Processing Agreement and source-code escrow, which follow company formation).

Control summary

"Evidenced", "Asserted" and "Roadmap" are used as defined at the top of this guide. Evidence can be shared on request under NDA.

Control area	What Rosa does	Status	Evidence
Data residency	UK/London (eu-west-2) only; no cross-region copy of customer data	Evidenced	Live configuration extract
Retention: access	7-day outputs and logs, 365-day model checkpoints; access refused on expiry	Evidenced	Configuration and behaviour test
Retention: deletion	Sweep removes artifact files, raw inputs, preserved failure logs and the per-job working directory; runs at most hourly while the instance is up, and once more at idle shutdown; a retired instance has its encrypted volume destroyed	**Evidenced (mechanism)**	Code and behaviour tests, 2026-08-19. Residual: an instance that is never started again does not sweep; removal on request at any time

Control area	What Rosa does	Status	Evidence
Integrity evidence	Run Manifest per job with SHA-256 input and output hashes	Evidenced	Manifest schema and live manifests
Off-site immutable evidence	Seven-year Object Lock (COMPLIANCE) copy of every manifest	Evidenced **where the vault is provisioned**	Shared instance vault provisioned; proven un-deletable against the live bucket
Encryption in transit (public)	TLS 1.2+/1.3, HSTS, auto-redirect	Evidenced	TLS configuration scan
Encryption in transit (internal)	Mutual TLS on every internal hop; per-cell CA with a non-exportable KMS-held signing key; automated rotation	Evidenced	Completed 9 Aug 2026
Encryption at rest	AES-256, AWS-managed keys, verified per volume	Evidenced	Per-volume verification
Customer-managed keys (BYOK)	Sole customer control of the key	Roadmap, dedicated	Not available today
Browser supply chain	Pinned self-hosted runtime under a restrictive CSP; egress limited to Rosa's API	Evidenced	Automated test asserts no third-party request
Browser injection-point guard	Build fails if a new dynamic-HTML injection point is added to the portal	Evidenced	Build guard in the release pipeline; see the two disclosed CSP limitations above
Portal security response headers	HSTS, frame, content-type and referrer headers	Roadmap	Not available on the portal's current static hosting; a hosting change is the remedy
Tenant isolation	Workspace-scoped access; authentication at the instance; negative-isolation tests covering reads, writes and the idempotency-key namespace	Evidenced as designed and tested	Test suite, plus an adversarial two-workspace test against the running shared service (6 September 2026, nine checks, nine refusals, all 404); independent test is Roadmap

Control area	What Rosa does	Status	Evidence
Dedicated cell isolation	Own cell: front door, hostname, network boundary, backup set	Offered, provisioned per engagement	Architecture supports it
Bring-your-own-cloud	Run inside your own cloud account	Planned, not available today	Roadmap
API authentication	argon2id verifier hashes; no plaintext key persisted server-side or logged; rotate and revoke	Evidenced	Key store schema and log inspection
Enterprise SSO / MFA / client mTLS	Federate to your identity provider	Roadmap	
Admin access control	Hardened per-instance secret; least-privilege cloud IAM; MFA on infrastructure change	Evidenced	IAM policy extract
Backups + tested restore	14 daily and 12 monthly snapshots; rehearsed dated restore	Evidenced for recoverability	14 Jul 2026, `integrity_check=ok`; recovery time not measured
Recovery objectives (RTO/RPO)	Agreed targets and higher-availability options	Roadmap, dedicated	
Monitoring and alerting	Automated fleet alerts; independent external prober	Evidenced	Alert configuration
Rate limiting / upload guard	Per-key and per-source limits; oversize uploads rejected pre-write; a stopped instance started only for a recognised key, at most 20 times an hour per source address on the shared instance	Evidenced	Configuration and tests; start limit verified in the running front door 13 Sep 2026
Secure SDLC	Peer-reviewed PRs, CI, dependency and secret scanning	Evidenced	CI configuration

Control area	What Rosa does	Status	Evidence
Change control / versioning	Provenance-stamped; customer-serving instances pinned per instance and updated deliberately	Evidenced	Manifest records the served image
Audit logging	Structured, append-only, 365-day retention; the shipping buffer is bounded and discards oldest under sustained pressure	Evidenced on mechanism	Log configuration; 22 Aug 2026
Customer SIEM export	Stream events into your monitoring	Roadmap, dedicated	
SOC 2	Designed toward Security and Confidentiality	Roadmap	Not audited
Independent penetration test	Third-party report	Roadmap	Not performed
Contractual SLA / incident response	Committed response and restoration times	Roadmap	Per engagement, formalised on incorporation
DPA / GDPR programme	Processor DPA, privacy notice, RoPA, DSAR procedure	Roadmap	On incorporation, expected Q4 2026
Cyber insurance	Cover appropriate to engagement size	Roadmap	On incorporation, expected Q4 2026
Delivery, succession and key-person cover	Founder-led operation today; specialist support engaged in security, legal and data protection, data science and engineering; security owner and guide approver currently the same person	Asserted; succession, key-person cover and role separation are Roadmap	Formalised alongside incorporation, expected Q4 2026; see Vendor and operating model

Vendor and operating model

Rosa is operated from the United Kingdom as a highly automated platform. **A UK limited company is being formed for the business.** On incorporation the contractual layer is put in place with it: a processor Data Processing Agreement, a privacy notice, a Record of Processing Activities and a DSAR procedure, a committed incident-response and breach-notification window, cyber insurance appropriate to engagement size, and business-continuity arrangements. Executed documents are available to counterparties on request shortly after formation.

Who delivers and operates Rosa today, stated plainly. Rosa is led and run day to day by its founder. Specialist support has been engaged in security, legal and data protection, data science and engineering, with contracts to follow, and further team members are due to join. **None of those additional people has started yet, so we do not describe a team in the present tense.** *Asserted.* Succession and key-person cover are being formalised alongside incorporation; they are not in place today. *Roadmap.* **The security owner and the approver of this guide are currently the same person.** Separating those two roles follows as the team forms, and this guide will say so when it has happened rather than before. We keep this disclosure here deliberately and update it as the position changes, because a disclosure that disappears from a document is harder to assess than one that stays and is qualified.

On contracting, today. There is no contractual service-level agreement and no committed breach-notification window yet; we will notify you promptly and in any event without undue delay, and we will agree a specific window in the Data Processing Agreement. Where an engagement needs a signed counterparty before those documents exist, tell us early - the sequencing is the thing to plan around, and we would rather say so than let it surface late in a procurement.

Subprocessors. Amazon Web Services (compute, storage, key management, all in eu-west-2) is the infrastructure subprocessor. A complete, current subprocessor list with the purpose of each is supplied on request and will be published with the Data Processing Agreement.

Reducing vendor dependency. A dedicated cell is the strongest answer available today: it is your own instance, with its own front door, network boundary and backup set, and its evidence is exported to object storage you can be given direct access to. **Bring-your-own-cloud - running Rosa inside your own AWS account - is planned for a future release and cannot be relied on in a decision made today.** Neither removes all dependency, since updates, support and fixes still come from us, and we would rather say that than overstate it. Source-code escrow is available to discuss for dedicated engagements.

Roadmap to enterprise and regulated readiness

Ahead of a first regulated engagement: a SOC 2 readiness assessment and an independent penetration test; customer-managed encryption keys on the dedicated tier; enterprise SSO and MFA; customer log export; defined recovery objectives with higher-availability options; evidence-vault provisioning as a standing prerequisite for every new instance; and, at company formation, the full contractual and privacy programme including the Data Processing Agreement, record of processing, insurance and a committed incident-response window.

Getting in touch

For a security questionnaire (CAIQ, SIG, or your own control set), a copy of a specific policy, evidence for any Evidenced item above, or a conversation about your requirements, contact **support@rosadebias.com**. We answer to the built state of the service, not the intended one.