

Nobody issues their own passport. We grade the model, sign the record, and publish either direction.

When a buyer, a board, or a regulator asks who graded a clinical AI model, the name on the record should not be the vendor's. RocSite is the independent adversarial validation layer: we do not build the model, we attack it on terms fixed before the data is opened, and we hand back a signed record that anyone can verify without trusting us. Offline. No self-grade. Receipts.

01 PRE-REGISTER

Hypothesis, plan and numeric thresholds locked and timestamped on a public registry before any data access.

02 PROBE, DO NOT BENCHMARK

Adversarial checks against locked clinical instruments: label instability, confounding, ablation, synthetic discrimination.

03 RECORD EVERYTHING

Every run emits a manifest: seed, versions, platform, per-phase results, warnings, summary.

04 ATTEST AND PUBLISH

Ed25519 signed by a named signer. Patient data never leaves its host. Results published whichever way they fall.

286,510

The public proof. Pre-registered falsification of sepsis prediction models across four independent ICU datasets (MIMIC-IV, MIMIC-III, eICU-CRD, PhysioNet Challenge 2019). Thresholds locked on OSF, 11 March 2026. Pooled, 4 of 16 pre-registered tests confirmed and the care-process shortcut hypothesis was not supported; we published that. Stratified by hospital size, the dependence crossed the locked threshold in hospitals under 100 beds, a post hoc finding we labelled as such. medRxiv 2026.03.17.26348414v2 · osf.io/9tbjm · code Apache-2.0.

WHAT IS RUNNING NOW

PRODUCT

WHAT IT DOES

Passport

One model, one named event: a Definition-ID, a PROvable or ATTESTED record, and the binding. RocSite signs; the vendor does not. Pilot 14 days, single-model passport 21 days, annual recert 10 days.

AI Governor v1, 21 modalities

Grades a model's claim against locked, peer-reviewed clinical scoring instruments and signs the receipt: input hash, instrument version, model output, deterministic output, divergence, timestamp. Not a medical device.

ICH Triage

Intracranial haemorrhage detection on non-contrast head CT, on site, offline, no PHI transmitted. FDA Class II SaMD pathway, Pre-Submission Q253397 accepted January 2026. Preliminary internal validation AUC 0.9592.

Discovery · Caliper

Pre-registered research design, and drug-safety claims graded against FAERS, with a definition stapled to the record.

WHO IT IS FOR

- **Hospitals, CMIOs, procurement.** You are being handed validation written by the model's builder on a benchmark they chose. Get the other side of the table, and a record you can staple to the decision.
- **AI vendors on an FDA pathway.** Independent adversarial evidence a regulator or procurement committee will accept, produced by someone who tried to break your model and could not.
- **Clinicians and researchers.** A pre-registered, falsifiable procedure you can point at your own question, with the code public and the result published either way.

TRUST POSTURE

- Offline by design; validation runs where your data already lives. No PHI leaves your environment.
- AES-256 at rest and in transit. Ed25519 signed, immutable evidence chain, independently verifiable.
- HIPAA-aligned; BAA available where a deployment handles PHI. SOC 2 Type II in progress.
- Every public claim carries a grade. Find one that does not and we correct it.

PRE-REGISTERED · FULL CODE PUBLIC · NULL RESULTS PUBLISHED · OFFLINE · CRYPTOGRAPHIC ATTESTATION

If your AI has to be right, it should first survive an attempt to prove it wrong. Bring us one model and one named event, and we will tell you the scope and grade path on the thread.

adam@rocsite.com

rocsite.com/sorbonne

rocsite.com/passport · rocsite.com/ai-governor

rocsitediscovery.com