Vaelon

An independent runtime-assurance (RTA) “safety cage” for autonomous satellite AI.

As spacecraft increasingly run onboard AI that makes its own decisions, there is no standard way to **prove a decision is safe before it reaches the bus**. Vaelon is an independent layer that sits between the onboard AI and the spacecraft: for every proposed action it returns **ALLOW / BLOCK / ESCALATE**, substitutes a verified safe fallback when needed, and cryptographically signs each decision into a tamper-evident record. Model-agnostic, it works over any onboard AI, and never trusts it, re-checking the safety-critical physics itself.

FORMALLY VERIFIED, AND BUILT TO FLY

- **Formally proven, not only tested**, 9 machine-checked proofs (Kani) hold over the **entire** input domain: it always fails closed on bad telemetry, can never approve an action that violates a hard safety invariant, and **stale ground telemetry can only tighten the gate, never loosen it**.
- **Grounded in real physics**, its orbital propagation matches the reference SGP4 to 1 metre; verdict-identical to an independent reference engine across 20,000 scenarios.
- **Flight-suitable engine**, deterministic Rust, ~5.6 KB base gate (≈26 KB full multi-mode engine), zero heap, zero `.data`/`.bss`, 14 ns/decision (42 ns p99); cross-compiles to ARM Cortex-M & RISC-V, links into C flight software (cFS / F′).
- **Signed & independently verifiable**, every decision folded into an HMAC-SHA256 chain a second, separate implementation reproduces byte-for-byte. Operators bound **their own** limits (incl. collision probability) as data, no recompile.

IT HOLDS AGAINST A REAL TRAINED AI, AT SCALE

100% → 0
of 3,000 randomized missions, a capable-but-unsafe neural policy breached a hard limit; with Vaelon guarding, power and fault breaches go to zero

0 / 2,000
breaches from an AI trained specifically to defeat the guard, across 2,000 unseen missions

> The measured result: power and fault breaches driven to zero across 3,000 missions, and collision avoided in 99.4% of conjunctions against independently-validated orbital dynamics (given a ground miss-distance and an along-track burn); the red-team also surfaced two real bugs, since fixed. Every number is reproducible, and the signed evidence is live: `vaelonspace.com/evidence`

WHERE IT FITS

> Vaelon **complements** conjunction-assessment and space-traffic-management tools, it doesn't replace them. Its edge is being **space-native** (it re-computes the safety-critical physics itself onboard rather than generic boundary-checking), **provenance-first** (every decision signed and independently verifiable), and **formally verifiable**, the concrete runtime independent-verification the ECSS “safety cage” names but has not yet specified.

HONEST STATUS

> An early-stage prototype (TRL 3-4), validated in software on real orbital data, with no flight heritage yet. That is precisely what incubation is for: an independent verification-and- validation step and a first in-orbit demonstration with a design partner.

THE ASK

ESA BIC UK incubation, plus an introduction to a **satellite operator or prime** running onboard autonomy for a **shadow-mode pilot**, they run Vaelon on their own mission data (no flight integration, an afternoon) and get a signed assurance report. That unlocks the two things left, independent V&V and a first flight, turning a validated prototype into a certifiable product. The market is early; that is the opportunity.