

Add to your application the control and the proof of every answer.

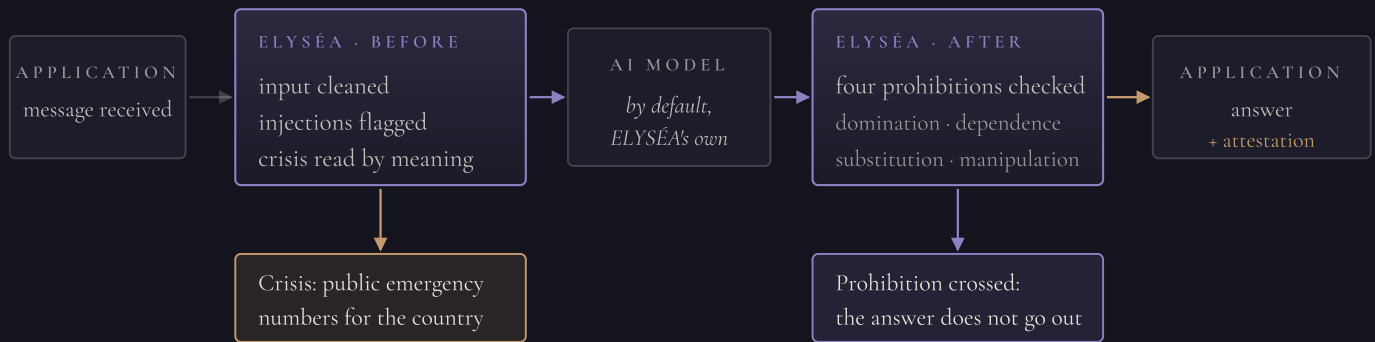
ELYSÉA is a software layer that you embed through an integration kit (SDK) between your application and the AI model: it checks every exchange before and after the model, and leaves an attestation on every answer.

YOUR PROBLEM

- **Prove, not just claim.** Model publishers now release dated safety cards, version by version; by building on top of them, you inherit that expectation without having the tools.^{1,2}
- **The insurers of your clients see nothing.** They exclude generative AI from their policies because they do not know which models their insureds run, nor for what use.³
- **A guardrail is paid for on every request.** It is judged on three axes at once: safety, wrongful refusals and added latency.⁴
- **The law now targets behaviour.** In the EU, telling people they are interacting with an AI and marking generated content; in California, a published crisis protocol and annual reports from 2027-07-01 for certain conversational applications.^{5,6}

HOW ELYSÉA ANSWERS IT

The user talks to the application, never to ELYSÉA



Your application receives the message; ELYSÉA checks it before the model and checks the answer after, and your application returns it with its attestation.

WHAT YOU GET	IN PRACTICE
Control before the model	Input cleaned; attempts to hijack the model's instructions (instruction injection) flagged before they reach it.
Control after the model	Four prohibitions checked on every answer: domination, dependence, substitution, manipulation. An answer that crosses one does not go out.
Rules that cannot be switched off	Your configuration can tighten the rules, never weaken them: any attempt is refused.
Crisis read by meaning	Distress is decided by the meaning of the message, not by a word list; the answer directs to public emergency numbers for 14 countries.
A per-answer attestation	Every answer carries an attestation tied to the exact version of the software that produced it.
Emotions	Emotion recognition locked in the layer, beyond the reach of configuration; a voice analysis path remains on the deployed service, removed with voice in the upcoming version.
The model	Today, ELYSÉA's model, which the upcoming version calls in Paris; on the deployed service, the provider setting is not verified; direct access to the model, possibly outside the EU, remains open by configuration. Mistral, OpenAI, Gemini and Claude are validated for use with your own model key, planned, not yet enabled.

VERIFIABLE BEFORE ANY COMMITMENT

- **Public:** the comparator demonstration at elysea.app/comparateur, the model alone versus the model with ELYSÉA, turn by turn (illustrative data, labelled); the kit published on npm as [@elysea/core](https://npmjs.com/package/@elysea/core).
- **On request, in a technical review:** the architecture, the control before and after the model, the prohibitions, the attestation, shown in the code.
- **On request:** the measurement protocol, written and sealed on 2026-08-28 before any measurement, amended since: four configurations compared, three judges from three model families, wrongful refusals measured alongside safety.
- **Decided on 2026-10-01, to be built:** a public benchmark where the annotations will be published and re-annotation opened to any laboratory.

The figures for safety, wrongful refusals and added latency will be published after the campaign, favourable or not, with their uncertainty. No figure is given before, and no third-party validation exists today: that is the condition for a measurement that does not adjust to its result. Next proofs, in order: upcoming version deployed and verified, pilot, new sealing and publication of the protocol, campaign, results published and re-annotatable by an independent laboratory in a public benchmark still to be built, entity registered.

WHAT SETS ELYSÉA APART

EXISTING APPROACH	WHAT IT DOES, ACCORDING TO ITS SOURCE	WHERE ELYSÉA DIFFERS
Content classification models: Llama Guard 3, ShieldGemma, OpenAI moderation	Classify a text by categories of risky content.	ELYSÉA also checks how the answer treats the person, and directs to help in a crisis. The two can be combined. ^{7,8,9}
Programmable guardrails: NeMo Guardrails, Bedrock Guardrails	Controls placed before and after the model, configured by the team itself.	Same position; ELYSÉA adds a floor that configuration cannot lower, and a per-answer attestation. ^{10,11}
Agent security: Lakera (Check Point group)	Inspects what goes in and what comes out, governs agent actions.	Centred on attacks; ELYSÉA is centred on the relationship with the person and on proof. ¹²

OFFER AND COMPLIANCE

Offer

- **Test:** 500 interactions per month, free of charge.
- **Production,** no subscription, per user active in the month: €2.50 excl. VAT up to 5,000, then €1.50 up to 25,000, then €0.80, each band at its own rate.
- **Sales:** open once the entity is registered; until then, test and pilot with no commitment.

Compliance

- **AI Act, information:** your application tells people they are interacting with an AI and marks generated content, from its placing on the market; ELYSÉA does not speak in its place.^{5,13}
- **AI Act, emotions:** inferring them at work and at school has been prohibited since 2025-02-02; the ELYSÉA layer locks it.⁵
- **Data:** database in Paris, model in Paris, accounts in Frankfurt. The upcoming version removes several non-EU paths from the core; other providers (website, email, payment, audience measurement, web search) remain with an unproven region, declared as such.
- **Contract:** data processing agreement in draft, signed when sales open.

NEXT STEP

A 45-minute technical demonstration: the architecture, the protocol, the comparator. Then a test key to integrate ELYSÉA into your own environment. contact@elysea.app · elysea.app/sdk

SOURCES (web pages read on 2026-10-01 unless dated) 1 Anthropic, report of 2026-08-17 — Transparency Hub 2 OpenAI, cards up to 2026-09-29 — Deployment Safety Hub 3 CSIS, G. C. Allen, 2026-09-04 — The Insurance Industry's Retreat from AI Threatens to Slow... 4 Artificial Analysis, 2026-06-11 — Benchmarking guardrail models for safety, refusal, and latency 5 EU — Regulation 2024/1689, OJ L of 2024-07-12 6 California Legislature — California SB 243 7 Meta — Llama Guard 3 — model card 8 Google — ShieldGemma — model card 9 OpenAI — Moderation — developer guide 10 NVIDIA — NeMo Guardrails — documentation overview 11 AWS — Amazon Bedrock Guardrails 12 Lakera, Check Point group — AI Agent Security 13 EU — Regulation 2026/1744, OJ L of 2026-07-24