
Dossier 01 — Builders and developers

Audience: teams building an AI application (consumer product, internal tool, platform) who want to integrate an ethical layer into it without changing model.

Identity: ELYSÉA is an ethical cognitive infrastructure embeddable through an SDK, a generalist brain. It does not talk to the user: the builder's application is what talks. ELYSÉA is neither a companion, nor a chatbot, nor a coach, nor a therapist, nor a wellness product, nor a medical device.

Status as of 2026-10-01: sales closed until the entity is registered (`elysea-web:src/lib/publicFacts.ts:69@00c9c89`). No campaign results: every performance figure is *to be measured* — *campaign in progress*.

I. Their problem

- **They must prove the safety of what they ship, not merely assert it.** Model publishers now release dated, per-version cards with their safety evaluations [S-PLT-05] [S-PLT-06]. A builder who builds on top of a model inherits that expectation without having the means for it.
- **Their customers and their customers' insurers ask for traces.** Insurers exclude generative AI from their policies because they cannot see which models their insureds are running, nor for which applications: “Carriers cannot see which models their insureds are running and for what applications” [S-ASS-01].
- **A guardrail costs something on every request.** “A guardrail is a tax on every request”: a guardrail is judged on three axes, safety, over-refusal and latency [S-MES-11].
- **Laws now target conversational behaviour.** In California, SB 243 requires operators of relational applications to publish a crisis protocol and to file annual reports [S-EDU-05]. In the EU, art. 50 of the AI Act requires informing the person that they are interacting with an AI [S-DRT-01].

2. What they must understand, prove and use (state of the art)

They must...	What the state of the art says	Source
Understand that a measurement is read together with its uncertainty	Statistics “reported with estimated uncertainties”; paired differences	[S-MES-01] [S-MES-10]
Require judges validated against humans	Compare with human annotation, several judges, inter-judge agreement	[S-MES-01] [S-MES-08]
Require rates corrected for judge bias	Correction by measured sensitivity and specificity	[S-MES-09]
Look at over-refusal as much as at safety	Specificity measured alongside F1 and latency	[S-MES-11]
Expect platform standards from an infrastructure	Versioned API with no unannounced breaking change, changelog, exportable audit logs, written responsibilities, vulnerability reporting channel	[S-PLT-01] [S-PLT-02] [S-PLT-04] [S-PLT-07]

3. Frequent objections and answers

Objection	Answer	Evidence
<i>I am locked into my model.</i>	The brain routes to four families of providers.	<code>elysea-core:functions/src/core/byok/-byokAllowList.ts:10-15@a5b9780e</code>
<i>I want my own key.</i>	The code that resolves the builder's key exists. Its activation is not wired: BYOK is not available today and will only be announced once proven on the deployed service.	<code>elysea-core:functions/src/core/byok/-byokKeyService.ts:2@a5b9780e · elysea-web:src/lib/publicFacts.ts:66-67@-00c9c89</code>
<i>Will a customer be able to switch safety off?</i>	No. A builder can tighten, never weaken: any attempt receives a 400 refusal.	<code>elysea-core:functions/src/endpoints/-sdkSurface.ts:139@a5b9780e · elysea-sdk:GUIDE_INTEGRATION_CONSTRUCTEUR.-md:13@5ff4916</code>
<i>It will block everything.</i>	Over-refusal is measured in the campaign (fp70 bench: ELYSÉA must not be more blocking than the bare model). Result: to be measured — campaign in progress.	<code>elysea-tests-e2e:audits/PROTOCOLE_CAMPAGNE_-4BRAS_3TIIRS_SCELLE_2026-08-28.md:323@c9898dd3</code>
<i>How much latency is added?</i>	To be measured — campaign in progress. No figure is given before then.	—
<i>Your tests are run by you.</i>	Yes. No third-party validation exists today. The mechanical judges and the 3-judge panel are validated on held-out sets (amendment R-70, status proposed, e2e branch #517: <code>elysea-tests-e2e:audits/AMENDEMENT_R70_-SENS_DECOUPE_STRUCTURELLE_2026-10-01.md:46@-67316397</code>); judge/human validation D23 is provided by Sébastien's annotation after the pilot. The protocol was preregistered and sealed on 08-28, then amended; resealing is a step still to come, and 9 changes made after sealing are awaiting their decision record (<code>holding-seb-docs:dossiers/RECONCILIATION_-PIECES_CAMPAGNE_2026-09-30.md:6@8560b37</code>). The public re-annotatable bench has been decided (10-01), not yet built.	<code>elysea-tests-e2e:audits/PROTOCOLE_CAMPAGNE_-4BRAS_3TIIRS_SCELLE_2026-08-28.md:3@c9898dd3</code>

4. What ELYSÉA does for them, with the evidence

Function	Evidence	State
Control before the model (sanitised input, OWASP LLMOI guard)	<code>elysea-core:functions/src/core/security/-owaspInputGuard.ts:5@a5b9780e</code>	On main
Control after the model (Do limits checked on the response)	<code>elysea-core:functions/src/core/pipeline/-validateLLMResponse.ts:1094@a5b9780e</code>	On main
The 4 prohibitions (domination, dependence, substitution, manipulation) built into the architecture, violation = response blocked	<code>elysea-core:functions/src/core/guardian/-checks/d0LimitCheck.ts:10-11@a5b9780e</code>	On main
Crisis decided by meaning (graded LLM vote), not by keywords	<code>elysea-core:functions/src/core/pipeline/-createPipeline.ts:1192@a5b9780e</code>	On main

Public emergency resources for 14 countries, 11 types of emergency	<code>elysea-core:functions/src/core/safety/-emergencyDirectory.ts:17-29@a5b9780e</code>	On main. No partnership with any crisis line.
Per-response attestation, with the build fingerprint	<code>elysea-core:functions/src/core/attestation/-interactionAttestor.ts:4@a5b9780e</code> · <code>elysea-core:functions/src/core/sdk/-sdkChatHandler.ts:160-161@a5b9780e</code>	On main
Emotion recognition locked off, not modifiable by the builder	<code>elysea-core:functions/src/core/sdk/-sdkAppService.ts:68@a5b9780e</code>	On main
SDK @elysea/core	<code>elysea-sdk:packages/core/package.json:2-3@-5ff4916</code>	Repository at 0.9.0; npm publishes 0.8.0

Performance of each function: to be measured — campaign in progress (results file: `holding-seb-docs:depot-zenodo-2026-09-05/RESULTATS_CAMPAGNE.md:2@8560b37`).

5. Compliance that concerns them

- **AI Act art. 50.1 and 50.2:** inform the person that they are talking to an AI; mark outputs in a machine-readable format. Applicable from placing on the market; the postponement to 2026-12-02 only covers systems already on the market before 2026-08-02 [S-DRT-01] [S-DRT-02]. **Decision:** the builder displays the AI notice from the first version; ELYSÉA cannot do it in their place since it does not talk.
- **AI Act art. 5(1)(f):** inferring emotions at work or at school is prohibited [S-DRT-01]. Lock P-01 removes this risk from the ELYSÉA layer.
- **GDPR:** roles to be qualified case by case [S-DRT-07]. Draft art. 28 DPA and draft art. 26 agreement (cross-application memory): unsigned drafts (`holding-seb-docs:juridique/DPA_CONSTRUCTEURS_-2026-09-10.md:1@8560b37`, `holding-seb-docs:juridique/ACCORD_RESPONSABILITE_CONJOINTE_ART26_2026-09-08.md:1@-8560b37`). **Decision:** no customer production release without a signed DPA.
- **SB 243** (if the builder's application falls within its definition): published crisis protocol, referral to a crisis line, reports from 2027-07-01 [S-EDU-05]. Directory P-07 provides the public numbers; the published protocol remains the builder's responsibility.

6. Integration path

1. Read the integration guide (`elysea-sdk:GUIDE_INTEGRATION_CONSTRUCTEUR.md:13@5ff4916`).
2. Test key `sk_test_`: 500 interactions per month (`elysea-web:src/lib/publicFacts.ts:36@00c9c89`).
3. Plug in the SDK, read the attestation returned with each response (P-08).
4. Add their own restrictions (tighten), never remove any (P-04).
5. Display the AI notice (art. 50) and publish their crisis protocol if subject to SB 243.
6. Sign the DPA and, if cross-application memory is used, the art. 26 agreement — when sales open.
7. Switch to an `sk_live_` key.

7. Offer and pricing (brochures of 09-30, sales closed)

Offer	Price	Evidence
Test	500 interactions / month	elysea-web:src/lib/publicFacts.ts:36@00c9c89
Production, tiered, no fixed fee	€2.50 excl. VAT / MAU up to 5,000; €1.50 from 5,001 to 25,000; €0.80 above	elysea-web:src/lib/publicFacts.ts:44-56@-00c9c89

8. Concrete next step

Open a test account and integrate in a test environment. No paid commitment before the entity is registered.

Gaps to close for this audience (observed on 10-01)

- No security.txt and no disclosure policy (RFC 9116 requires Contact and Expires [S-PLT-07]) — no evidence on main.
- No public versioned API changelog in the sense of [S-PLT-02].
- No third-party audit report (SOC 2, ISO 27001) — the reference platforms have them [S-PLT-01] [S-PLT-03].

Dossier sources

S-PLT-01 · S-PLT-02 · S-PLT-03 · S-PLT-04 · S-PLT-05 · S-PLT-06 · S-PLT-07 · S-ASS-01 · S-MES-01 · S-MES-08 · S-MES-09 · S-MES-10 · S-MES-11 · S-DRT-01 · S-DRT-02 · S-DRT-07 · S-EDU-05