

Ajoutez à votre application le contrôle et la preuve de chaque réponse.

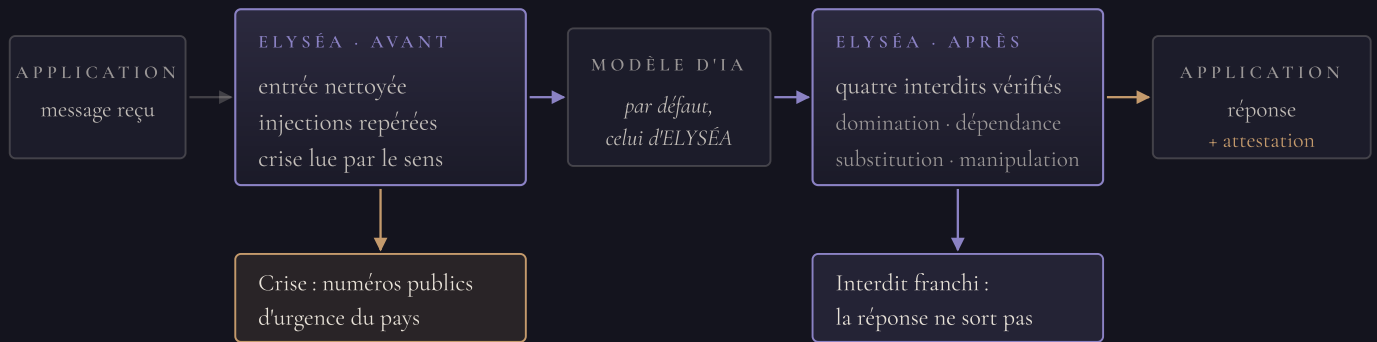
ELYSÉA est une couche logicielle que vous embarquez par un kit d'intégration (SDK) entre votre application et le modèle d'IA : elle contrôle chaque échange avant et après le modèle, et laisse une attestation sur chaque réponse.

VOTRE PROBLÈME

- **Prouver, pas seulement affirmer.** Les éditeurs de modèles publient désormais des fiches de sécurité datées, version par version ; en construisant au-dessus d'eux, vous héritez de cette attente sans en avoir les outils.^{1,2}
- **Les assureurs de vos clients ne voient rien.** Ils excluent l'IA générative de leurs polices parce qu'ils ne savent pas quels modèles tournent chez leurs assurés, ni pour quel usage.³
- **Un garde-fou se paie à chaque requête.** Il se juge sur trois axes à la fois : la sécurité, les refus à tort et la latence ajoutée.⁴
- **La loi vise le comportement.** Dans l'UE, informer la personne qu'elle échange avec une IA et marquer les contenus produits ; en Californie, un protocole de crise publié et des rapports annuels à partir du 01/07/2027 pour certaines applications conversationnelles.^{5,6}

COMMENT ELYSÉA Y RÉPOND

L'utilisateur parle à l'application, jamais à ELYSÉA



Votre application reçoit le message ; ELYSÉA le contrôle avant le modèle, contrôle la réponse après, et votre application la rend avec son attestation.

CE QUE VOUS OBTENEZ	CONCRÈTEMENT
Contrôle avant le modèle	Entrée nettoyée ; tentatives de détourner les consignes du modèle (injection d'instructions) repérées avant qu'elles ne l'atteignent.
Contrôle après le modèle	Quatre interdits vérifiés sur chaque réponse : domination, dépendance, substitution, manipulation. Une réponse qui en franchit un ne sort pas.
Règles non désactivables	Votre configuration peut durcir les règles, jamais les affaiblir : toute tentative est refusée.
Crise lue par le sens	La détresse est décidée par le sens du message, pas par une liste de mots ; la réponse oriente vers les numéros publics d'urgence de 14 pays.
Une attestation par réponse	Chaque réponse porte une attestation liée à la version exacte du logiciel qui l'a produite.
Émotions	Reconnaissance des émotions verrouillée dans la couche, hors de portée de la configuration ; un chemin d'analyse vocale subsiste sur le service déployé, retiré avec la voix dans la version en préparation.
Le modèle	Aujourd'hui, le modèle d'ELYSÉA, que la version en préparation appelle à Paris ; sur le service déployé, le réglage du fournisseur n'est pas vérifié ; un accès direct au modèle, hors UE possible, reste ouvert par configuration. Mistral, OpenAI, Gemini et Claude sont validés pour l'usage de votre propre clé de modèle, prévu, pas encore activé.

VÉRIFIABLE AVANT TOUT ENGAGEMENT

- **Public** : la démonstration du comparateur sur elysea.app/comparateur, le modèle seul face au modèle avec ELYSÉA, tour par tour (données illustratives, marquées) ; le kit publié sur npm sous le nom `@elysea/core`.
- **Sur demande, en revue technique** : l'architecture, le contrôle avant et après le modèle, les interdits, l'attestation, montrés dans le code.
- **Sur demande** : le protocole de mesure, écrit et scellé le 28/08/2026 avant toute mesure, amendé depuis : quatre configurations comparées, trois juges de trois familles de modèles, refus à tort mesurés à côté de la sécurité.
- **Décidé le 01/10/2026, à construire** : un banc public où les annotations seront publiées et la réannotation ouverte à tout laboratoire.

Les chiffres de sécurité, de refus à tort et de latence ajoutée seront publiés après la campagne, favorables ou non, avec leur incertitude. Aucun chiffre n'est donné avant, aucune validation tierce n'existe aujourd'hui : c'est la condition d'une mesure qui ne s'ajuste pas au résultat. Prochaines preuves, dans l'ordre : version en préparation déployée et vérifiée, pilote, nouveau scellement et publication du protocole, campagne, résultats publiés et réannotables par un laboratoire indépendant dans un banc public à construire, entité inscrite.

CE QUI DISTINGUE ELYSÉA

APPROCHE EXISTANTE	CE QU'ELLE FAIT, SELON SA SOURCE	OÙ ELYSÉA DIFFÈRE
Modèles de classement de contenu : Llama Guard 3, ShieldGemma, modération OpenAI	Classent un texte par catégories de contenu à risque.	ELYSÉA vérifie aussi la conduite de la réponse envers la personne, et oriente en cas de crise. Les deux se combinent. ^{7,8,9}
Garde-fous programmables : NeMo Guardrails, Bedrock Guardrails	Des contrôles placés avant et après le modèle, que l'équipe configure elle-même.	Même position ; ELYSÉA ajoute un socle que la configuration ne peut pas abaisser, et une attestation par réponse. ^{10,11}
Sécurité des agents : Lakera (groupe Check Point)	Inspecte ce qui entre et ce qui sort, gouverne les actions des agents.	Centrée sur l'attaque ; ELYSÉA est centrée sur la relation avec la personne et sur la preuve. ¹²

OFFRE ET CONFORMITÉ

Offre

- **Test** : 500 interactions par mois, sans frais.
- **Production**, sans abonnement, par utilisateur actif dans le mois : 2,50 € HT jusqu'à 5 000, puis 1,50 € jusqu'à 25 000, puis 0,80 €, chaque tranche à son tarif.
- **Ventes** : ouvertes à l'inscription de l'entité ; d'ici là, test et pilote sans engagement.

Conformité

- **AI Act, information** : votre application informe la personne qu'elle échange avec une IA et marque les contenus produits, dès sa mise sur le marché ; ELYSÉA ne parle pas à sa place.^{5,13}
- **AI Act, émotions** : leur inférence est interdite au travail et à l'école depuis le 02/02/2025 ; la couche ELYSÉA la verrouille.⁵
- **Données** : base à Paris, modèle à Paris, comptes à Francfort. La version en préparation retire du cœur plusieurs chemins hors UE ; d'autres prestataires (site, courriels, paiement, mesure d'audience, recherche web) restent à région non prouvée, déclarés comme tels.
- **Contrat** : accord de sous-traitance en brouillon, signé à l'ouverture des ventes.

PROCHAINE ÉTAPE

Une démonstration technique de 45 minutes : l'architecture, le protocole, le comparateur. Puis une clé de test pour intégrer ELYSÉA dans votre propre environnement. contact@elysea.app · elysea.app/sdk

SOURCES (pages web lues le 01/10/2026, sauf date indiquée) 1 Anthropic, rapport du 17/08/2026 — Transparency Hub 2 OpenAI, fiches jusqu'au 29/09/2026 — Deployment Safety Hub 3 CSIS, G. C. Allen, 04/09/2026 — The Insurance Industry's Retreat from AI Threatens to Slow... 4 Artificial Analysis, 11/06/2026 — Benchmarking guardrail models for safety, refusal, and latency 5 UE — Règlement 2024/1689, JO L du 12/07/2024 6 California Legislature — California SB 243 7 Meta — Llama Guard 3 — fiche du modèle 8 Google — ShieldGemma — fiche du modèle 9 OpenAI — Moderation — guide développeur 10 NVIDIA — NeMo Guardrails — documentation, vue d'ensemble 11 AWS — Amazon Bedrock Guardrails 12 Lakera, groupe Check Point — AI Agent Security 13 UE — Règlement 2026/1744, JO L du 24/07/2026