Intelligence artificielle

Version du 19 novembre 2026

Ce que fait WYRLO avec l'intelligence artificielle, ce qu'il ne garantit pas, et ce que vous devez dire à vos propres utilisateurs. Page d'information au titre du règlement (UE) 2024/1689 sur l'intelligence artificielle ; les engagements contractuels figurent dans les CGU.

Sommaire

- Ce que fait WYRLO
- Modèles utilisés
- Ce que WYRLO ne garantit pas
- Règlement européen sur l'IA
- Ce que vous devez faire vis-à-vis de vos utilisateurs
- Questions et signalements

Ce que fait WYRLO

- Vérification : pour chaque réponse de votre RAG, WYRLO repère les injections de prompt par des règles, compare la réponse aux passages fournis et, en français et en anglais, fait vérifier par un modèle de langue de Mistral AI que chaque affirmation figure dans les passages. Il en tire un niveau de confiance et une décision : répondre, ou dire « je ne sais pas ».
- Index sans copie : un modèle de vecteurs de Mistral AI transforme les passages en représentations numériques, pour la recherche. Le texte n'est pas gardé.
- Rédaction (/v1/answer, en préversion) : un modèle de langue rédige une réponse citée à partir de vos passages, puis WYRLO la

vérifie. Le modèle est celui géré par WYRLO (Mistral AI) ou celui que vous choisissez.

- Amélioration de nuit : à partir des traces de chaque projet, WYRLO propose et teste des réglages (vocabulaire, passages gardés, seuils de confiance). Un changement n'est adopté que s'il améliore les décisions mesurées sans en dégrader ; il est surveillé ensuite et annulé automatiquement s'il dégrade le résultat. Un modèle de langue rédige l'explication des échecs et le rapport.
- Assistant d'aide : dans l'espace développeurs, un modèle de langue répond à vos questions sur WYRLO à partir du centre d'aide, et sa réponse est vérifiée avant d'être affichée.

Modèles utilisés

- Mistral AI (France), par son point d'accès européen : modèles de langue (mistral-small, mistral-large) et modèle de vecteurs (mistral-embed). Ils ne reçoivent que des textes masqués par WYRLO.
- Le modèle de votre choix, pour /v1/answer : OpenAI, Anthropic, Mistral AI, Azure OpenAI ou une adresse compatible, appelé avec votre clé, sans masquage.
- Des règles écrites par WYRLO (injections, masquage, contrôles de cohérence), qui décident seules quand un modèle n'est pas disponible.
- WYRLO n'entraîne aucun modèle avec vos données, ni les siens ni ceux d'un tiers.

Ce que WYRLO ne garantit pas

- L'exactitude d'une réponse : une réponse marquée « certaine » peut être fausse, et un « je ne sais pas » peut écarter une bonne réponse. Le niveau de confiance est une estimation.

- La vérification ne juge que l'accord entre la réponse et les passages fournis : si vos documents sont faux, périmés ou incomplets, une réponse fidèle à ces documents le sera aussi.
- La vérification par un modèle de langue n'existe qu'en français et en anglais ; dans les autres langues, le niveau de confiance repose sur d'autres signaux, moins fiables.
- Quand WYRLO ne répond pas à temps, que le volume de l'offre est atteint ou qu'un modèle est indisponible, la réponse passe marquée « non vérifiée ».
- Le repérage des injections et le masquage sont automatiques et ne sont pas exhaustifs ; les noms de personnes ne sont pas masqués.

Règlement européen sur l'IA

WYRLO fournit un service d'IA à des professionnels, qui l'intègrent dans leurs propres applications. Il utilise des modèles d'IA à usage général de Mistral AI, qui en est le fournisseur. WYRLO n'est pas conçu pour les usages à haut risque énumérés à l'annexe III du règlement (emploi, éducation, crédit, services publics essentiels, maintien de l'ordre, justice, migrations, etc.) et ne prend aucune décision sur des personnes : il décide seulement si une réponse peut être montrée.

Les obligations de transparence de l'article 50 du règlement envers les personnes qui échangent avec une IA, ou lisent un texte produit par une IA, reposent d'abord sur l'organisation qui met l'application à leur disposition, c'est-à-dire vous.

Ce que vous devez faire vis-à-vis de vos utilisateurs

- Les informer, de façon claire et au plus tard lors de leur première interaction, qu'ils échangent avec un système d'IA, sauf si c'est

évident ; et qu'une réponse rédigée par /v1/answer est produite par une IA.
- Leur expliquer ce que signifient « je ne sais pas » et « non vérifiée », et leur offrir un recours humain quand l'enjeu le justifie.
- Ne pas utiliser WYRLO comme seul fondement d'une décision produisant des effets juridiques sur une personne, ni dans un système interdit par l'article 5 du règlement ; nous prévenir avant toute intégration dans un système à haut risque.
- Former les personnes de votre organisation qui utilisent ou intègrent WYRLO à ses limites (maîtrise de l'IA, article 4 du règlement).
- Informer vos utilisateurs du traitement de leurs données (questions, réponses) et du recours à WYRLO et à ses sous-traitants, et ne pas envoyer de données sensibles sans les masquer.
- Si vous êtes Partenaire, transmettre ces obligations à vos Clients finaux.

Questions et signalements

Pour toute question sur l'IA dans WYRLO, ou pour signaler une réponse problématique ou un risque, écrivez à contact@wyrlo.io.

Conditions générales d'utilisation Sécurité et données

Sous-traitants