Notice d'information

Version 1.0 · Entrée en vigueur : 15 septembre 2026

Cette notice décrit le fonctionnement du système d'intelligence artificielle mis en œuvre par REPERES, ses limites et les précautions d'usage. Elle répond à l'obligation de transparence applicable aux systèmes d'IA destinés à interagir avec des personnes physiques.

1. Vous interagissez avec un système d'intelligence artificielle

Les réponses produites par REPERES sont générées automatiquement par des modèles de langage. Aucun professionnel de BMSE ne relit une réponse avant qu'elle ne vous soit présentée.

2. Comment une réponse est produite

Une requête traverse cinq étapes.

1. Qualification. Un modèle de langage analyse votre question pour en déterminer l'intention, les thèmes et les secteurs concernés, et pour en enrichir la formulation par des mots-clés. Votre question n'est pas reformulée : les mots-clés s'y ajoutent, afin de ne pas écraser la spécificité de votre demande.

2. Recherche documentaire. La requête enrichie est transformée en représentation numérique, puis confrontée au corpus selon deux méthodes combinées, l'une fondée sur le sens et l'autre sur les termes exacts. Cette double approche permet de retrouver aussi bien une référence précise qu'une situation décrite en langage courant.

3. Classement. Un modèle de langage examine les extraits retenus et les classe selon leur pertinence pour votre question, en cinq niveaux : fondamental, utile, périphérique, hors sujet, non applicable au contexte de la requête. Les extraits hors sujet et non applicables sont écartés.

4. Rédaction. Un modèle de langage rédige la réponse à partir des extraits retenus et de la doctrine de restitution propre à REPERES. Selon la complexité de la demande, un modèle plus puissant est mobilisé.

5. Restitution des sources. Les documents mobilisés sont présentés à côté de la réponse. Chacun renvoie vers une page de référencement indiquant la source, l'éditeur, la date et le lien vers le document d'origine.

3. Les modèles employés

Le service met en œuvre plusieurs modèles de langage, chacun affecté à une étape précise du traitement. **L'intégralité de ces traitements est réalisée sur le territoire français**, chez des fournisseurs établis en France.

La liste nominative des modèles en service, de leurs fournisseurs et du lieu de leur traitement figure en **annexe, datée**. Elle est mise à jour à chaque changement, sans que la présente notice soit elle-même révisée.

Aucun de ces fournisseurs ne reçoit de données de la part de BMSE à des fins d'entraînement de ses modèles.

4. Le corpus

Le corpus est constitué exclusivement de sources publiquement accessibles : textes législatifs et réglementaires, jurisprudence publiée, conventions collectives étendues, recommandations et guides d'organismes publics, rapports publics.

Il ne contient aucun dossier individuel et aucun document interne d'organisme gestionnaire.

Le corpus reflète l'état des sources à la date de leur intégration. Une source peut avoir été modifiée ou abrogée depuis.

5. Ce que le système ne sait pas faire

Il peut se tromper. Les modèles de langage produisent parfois des affirmations inexactes, des rapprochements erronés ou des références imprécises, y compris lorsqu'ils s'appuient sur des sources correctes. Une réponse formulée avec assurance n'est pas pour autant exacte.

Il ne connaît pas votre situation. Le système ne dispose ni de votre dossier, ni de l'historique de votre établissement, ni des positions de votre autorité de tarification. Sa réponse est générale par construction.

Il ne remplace pas une vérification. Toute référence doit être vérifiée à sa source avant usage. Les liens vers les documents d'origine sont fournis à cette fin.

Il ne rend pas de décision. Une réponse n'est ni un avis, ni une décision, ni un acte opposable.

Il ne se souvient de rien. Vos questions ne sont pas conservées, et le système ne s'améliore pas à partir de vos usages.

6. La fonction de protection des données personnelles

6.1 Ce qu'elle fait

Lorsque vous l'activez, votre question et vos pièces jointes sont traitées avant transmission par LaPlume, un dispositif de reconnaissance d'entités nommées spécialisé sur le secteur social et médico-social. Les éléments identifiants détectés sont remplacés par des marqueurs neutres, qui préservent le sens et la structure du texte.

Le contenu protégé vous est présenté avant envoi. Vous le vérifiez et confirmez la transmission. Cette étape de prévisualisation est le moment où le risque résiduel décrit ci-dessous peut être corrigé.

Le dispositif combine un modèle de langue spécialisé, un moteur de règles pour les formats structurés, et des répertoires sectoriels. Sa taxonomie couvre 101 catégories, hiérarchisées en quatre niveaux de criticité, incluant des entités que les outils généralistes ne couvrent pas : types d'établissements, organismes publics du secteur, mesures de protection, pathologies du champ psycho-social.

Le traitement s'effectue en mémoire, en France. Ni le texte d'origine ni la correspondance entre éléments d'origine et marqueurs ne sont conservés. La réponse produite n'est pas ré-identifiée.

6.2 Ses performances mesurées

Le dispositif a fait l'objet d'une évaluation par annotation humaine indépendante sur dix rapports sociaux réels sélectionnés pour leur difficulté, avec trois annotateurs, plan croisé et arbitrage de 313 désaccords par un tiers.

NIVEAU DE CRITICITÉ	EXEMPLES D'ENTITÉS	SCORE F1
Identifiants directs	Noms, dates de naissance, numéros, téléphones, adresses électroniques	98,4 %
Quasi-identifiants forts	Adresses, pathologies, établissements et organismes nommés	97,3 %
Quasi-identifiants faibles	Organismes génériques, communes, départements	95,6 %
Contexte résiduel	Durées relatives, lieux génériques	83,5 %

Score global pondéré par criticité : **96,7 %**.

6.3 Ses limites connues

Le dispositif n'est pas une anonymisation. Un score de 98,4 % sur les identifiants directs signifie qu'une entité sur soixante environ échappe à la détection.

Certaines institutions locales ne sont pas couvertes. Les services non référencés dans les bases nationales ou nommés de façon non standardisée peuvent être manqués.

Un risque résiduel subsiste même après une protection réussie. Un récit peut rester reconnaissable par un professionnel du même territoire, sans qu'aucun nom n'y figure : la combinaison d'un type de structure, d'une géographie et d'un événement marquant peut suffire. Ce phénomène, distinct de la réidentification au sens strict, fait l'objet de travaux de recherche publiés par ConfidensIA. Il est plus marqué sur les situations singulières, sur les territoires restreints, et pour les lecteurs connaissant le réseau local.

C'est pourquoi les conditions générales d'utilisation vous interdisent de façon absolue de transmettre les identifiants directs d'une personne accompagnée et les pièces de dossier non expurgées, **que la fonction soit activée ou non**, et vous demandent de veiller à ce que la combinaison des éléments transmis ne rende pas une personne reconnaissable. La fonction et sa prévisualisation vous aident à respecter cette règle ; elles ne s'y substituent pas.

6.4 Travaux publiés

L'architecture du dispositif, son protocole d'évaluation et le modèle de risque associé sont documentés dans deux documents de travail publiés. Le modèle de reconnaissance d'entités et son corpus de distillation sont publiés sous DOI 10.5281/zenodo.17689395. Le pipeline complet demeure propriétaire.

7. Usage via un assistant conversationnel tiers

Lorsque vous interrogez REPERES depuis un assistant tiers, deux systèmes d'intelligence artificielle interviennent successivement : celui de votre assistant, puis celui de REPERES.

Votre question est d'abord traitée par le fournisseur de votre assistant, selon ses propres conditions, avant de parvenir à REPERES. Cette étape échappe au contrôle de BMSE. **La fonction de protection des données personnelles n'est pas applicable sur ce chemin.**

Votre assistant peut par ailleurs reformuler votre demande avant de la transmettre, et il peut reformuler la réponse de REPERES avant de vous la présenter. BMSE ne maîtrise ni l'une ni l'autre de ces transformations.

8. Signaler une erreur

Toute erreur constatée peut être signalée depuis le formulaire prévu dans le service. Les signalements portant sur une erreur de droit ou une source périmée sont traités en priorité et alimentent la correction du corpus.

9. Contact

Pour toute question relative au fonctionnement du système : reperes@bmse.fr.

Annexe — Modèles en service

État au 15 septembre 2026

ÉTAPE	MODÈLE	FOURNISSEUR	LIEU DU TRAITEMENT
Qualification	Mistral Small	Mistral AI	France
Recherche documentaire	Qwen3-Embedding-8B	Servi par Scaleway	France, `fr-par`
Classement des extraits	Mistral Small	Mistral AI	France
Rédaction, demandes courantes	Mistral Small	Mistral AI	France
Rédaction, demandes complexes	GLM 5.2, modèle à poids ouverts	Servi par Mistral AI	France
Lecture de documents joints	Mistral OCR	Mistral AI	France
Protection des données personnelles	LaPlume, modèle CamemBERT distillé	ConfidensIA SAS	France

Le modèle utilisé pour la rédaction des demandes complexes est un modèle à poids ouverts d'origine tierce, servi par Mistral AI sur son infrastructure située en France. Son auteur n'intervient pas dans la chaîne de traitement et ne reçoit aucune donnée.

Les versions antérieures de la présente annexe demeurent consultables.

Barbaux Morales Solutions et Expertise (BMSE) · SARL au capital de 1 000 € · SIREN 990 220 907 · RCS Saverne
3B rue des Jardiniers, 67190 Mutzig · reperes@bmse.fr · Protection des données : dpo@bmse.fr