

État de la confidentialité de l'IA 2026

Avocats, experts-comptables, services RH, banques et établissements de santé confient désormais des dossiers de clients et de salariés à Claude, ChatGPT, Gemini, Copilot et Le Chat. Ce rapport fait le point sur ce que la loi et les régulateurs leur demandent, sur ce que déclarent les outils audités par le CNB, sur la capacité des outils actuels à détecter les données personnelles françaises et marocaines, et sur ce qu'un cabinet peut faire dès cette année.

Hidden Tokens · 1^{re} édition · publié le 25 octobre 2026 · faits vérifiés les 24 et 25 octobre 2026

14/15

outils d'IA juridique de l'audit du CNB (2025) citent une société américaine parmi celles qui hébergent le LLM.

Grille CNB, juin 2025 ¹

9/15

ne déclarent rien sur la pseudonymisation ou l'anonymisation (« N/C » ou « N/A »).

Grille CNB, juin 2025 ¹

77,5 %

des identifiants directs français et marocains repérés par Presidio avec ses réglages par défaut, sur HT-Bench.

HT-Bench 0.2.0, 25 oct. 2026
¹¹

CONSTATS

- **Les règles existent déjà ; l'IA les rend urgentes.** Le secret professionnel, le RGPD et la loi marocaine 09-08 s'appliquent à un prompt comme à un courriel. L'AI Act ajoute des obligations propres à l'IA, sans remplacer aucune d'elles.
- **Le CNB nomme le problème.** Sa grille de lecture écarte, au titre de la souveraineté, l'hébergement par des sociétés soumises à des lois de réquisition extraterritoriales, et cite en exemple les sociétés américaines. Quatorze des quinze outils qu'elle résume citent une société américaine parmi celles qui hébergent le modèle de langage.
- **La pseudonymisation est la case vide.** Cinq fiches sur quinze seulement déclarent une pseudonymisation (3) ou une anonymisation (2). Le guide déontologique du CNB d'avril 2026 recommande de pseudonymiser ou d'anonymiser les données avant d'utiliser un outil d'IA générative.
- **Les détecteurs génériques ratent les identifiants locaux.** Sur 366 documents synthétiques français et marocains, Presidio par défaut a repéré 77,5 % des identifiants directs, dont 69,8 % des numéros de CIN, 5,5 % des RIB et 1,8 % des raisons sociales. Un outil conçu pour ces formats (Hidden Tokens 2.28.0) est à 99 % ou plus dans chaque pack sur HT-Bench (100 % sur le jeu de test 2.28.0). C'est un banc synthétique passé par l'éditeur de l'outil, pas un résultat sur de vrais dossiers.
- **Aucune catégorie ne règle tout.** Les offres entreprise des fournisseurs d'IA encadrent ce que le fournisseur fait de données qu'il reçoit en clair ; la DLP bloque ou retire ; les API de caviardage demandent un projet d'ingénierie. La pseudonymisation réversible sur le poste est la catégorie la plus récente et la moins mesurée, la nôtre comprise.

1 Pourquoi maintenant

En 2025, 20,0 % des entreprises de l'Union européenne d'au moins dix salariés utilisaient l'intelligence artificielle, contre 13,5 % un an plus tôt, et 8,8 % s'en servaient pour générer du langage écrit ou parlé³. En France, 18 % des entreprises utilisaient au moins une technologie d'IA en 2025, huit points de plus qu'en 2024 ; 43 % d'entre elles utilisaient des outils d'IA générative de texte ou de parole⁴. Parmi les outils que les équipes utilisent figurent les assistants généralistes : Claude, ChatGPT, Gemini, Microsoft Copilot et l'assistant de Mistral (Le Chat, renommé Vibe en août 2026).

Chacun de ces assistants est, par défaut, un service en ligne : le prompt et chaque fichier joint partent sur les serveurs du fournisseur, où le modèle les lit en clair. Les offres professionnelles ajoutent des engagements contractuels (pas d'entraînement sur les données du client par défaut, maîtrise de la conservation, parfois choix de la région d'hébergement), mais elles ne changent pas ce fait. Pour les professions tenues au secret, la question n'est pas de savoir si le fournisseur est digne de confiance. Elle est de savoir si le cabinet a le droit de lui remettre l'information.

Le secret professionnel

En France, le secret professionnel de l'avocat couvre, « en toutes matières », les consultations, les correspondances, les notes d'entretien et toutes les pièces du dossier (loi n° 71-1130 du 31 décembre

1971, article 66-5)⁹ ; la révélation d'une information à caractère secret par une personne qui en est dépositaire par profession est punie d'un an d'emprisonnement et de 15 000 € d'amende (Code pénal, article 226-13)¹⁰. Le Code pénal marocain réprime la même atteinte (article 446). Le secret appartient au client, non au cabinet : un cabinet ne peut pas y renoncer en acceptant les conditions d'utilisation d'un fournisseur.

Le RGPD

Un prompt qui contient des noms, des numéros d'identité ou les faits du dossier d'une personne identifiable est une donnée personnelle. L'envoyer à un fournisseur d'IA fait de celui-ci un sous-traitant ou un destinataire (article 28), et peut constituer un transfert hors de l'Union (chapitre V). Le RGPD définit la pseudonymisation comme un traitement après lequel les données « ne peuvent plus être attribuées à une personne concernée précise sans avoir recours à des informations supplémentaires » conservées séparément (article 4, 5) ; les données pseudonymisées restent des données personnelles pour qui peut les réidentifier (considérant 26), et les articles 25 et 32 citent la pseudonymisation comme mesure de protection⁵. Les lignes directrices 01/2025 du Comité européen de la protection des données (CEPD) sur la pseudonymisation ont été adoptées le 16 janvier 2025 en version soumise à consultation ; aucune version définitive n'était publiée en octobre 2026⁶. Dans l'arrêt *Contrôleur européen de la protection des données c. Conseil de résolution unique* (4 septembre 2025), la Cour de justice a jugé que des données pseudonymisées ne constituent pas des données personnelles « dans tous les cas et pour toute personne » : pour un destinataire qui ne peut raisonnablement pas les réidentifier, elles peuvent ne pas l'être, tandis que les obligations du responsable s'apprécient de son propre point de vue⁷. En pratique, la pseudonymisation ne fait pas sortir un cabinet du RGPD. Elle peut réduire ce que reçoit le fournisseur d'IA à des données qu'il ne peut attribuer à personne, à condition que la clé reste au cabinet.

Maroc : la loi 09-08

La loi n° 09-08 du 18 février 2009 encadre les traitements de données à caractère personnel au Maroc et a institué la Commission nationale de contrôle, la CNDP (article 27)⁸. Elle s'applique au traitement fait au moyen d'un outil d'IA comme à tout autre. L'article 43 ne permet de transférer des données vers un État étranger que si cet État assure un niveau de protection suffisant ; l'article 44 prévoit les dérogations, dont le consentement exprès de la personne et l'autorisation expresse et motivée de la CNDP⁸. La plupart des assistants d'IA généralistes traitent les prompts hors du Maroc : leur confier des données de clients pose d'abord la question du transfert. Le décret d'application demande déjà que les données transmises à un tiers à des fins statistiques ou de recherche soient « rendues anonymes ou codées » (décret n° 2-09-165, article 31)⁸. Nous n'avons trouvé aucune délibération de la CNDP propre à l'IA générative en 2025-2026, et le site de la CNDP n'a pas pu être consulté pour vérifier ses dernières publications.

Le calendrier de l'AI Act

Le règlement (UE) 2024/1689 est entré en vigueur le 1^{er} août 2024. Ses dispositions générales (dont l'obligation de maîtrise de l'IA) et ses pratiques interdites s'appliquent depuis le 2 février 2025, les règles sur les modèles d'IA à usage général depuis le 2 août 2025, et le règlement dans son ensemble depuis le 2 août 2026¹³. L'Omnibus numérique sur l'IA, règlement (UE) 2026/1744 (publié le 24 juillet 2026, en vigueur depuis le 27 juillet 2026), a reporté les obligations relatives aux systèmes à haut

risque : au 2 décembre 2027 pour ceux de l'annexe III et au 2 août 2028 pour ceux de l'annexe I¹⁴. L'annexe III vise notamment l'IA utilisée dans les décisions d'emploi, ce qui concerne les services RH. Pour un cabinet qui se contente d'utiliser un assistant généraliste sur des dossiers, la confidentialité reste régie par le secret et par le droit des données personnelles, pas par l'AI Act. Le volet RGPD de l'Omnibus numérique (COM(2025) 837), qui modifierait la définition de la donnée personnelle et ajouterait une disposition sur la pseudonymisation, était encore en négociation en octobre 2026¹⁵.

2 Ce que demandent les régulateurs : l'audit de 15 outils d'IA juridique par le CNB

Le texte le plus concret qu'un ordre professionnel ait publié sur le sujet vient du Conseil national des barreaux. La *Grille de lecture* de son groupe de travail sur l'intelligence artificielle (1^{re} édition, juin 2025, v1.2)¹ fait deux choses. Elle pose les critères de choix d'un outil d'IA juridique, et elle résume quinze outils consultés dans une fiche de synthèse de dix champs : lieu d'hébergement du LLM et société qui l'héberge, lieu d'hébergement des données et société qui les héberge, chiffrement, RAG sécurisée, réutilisation des données des clients par le LLM, pseudonymisation ou anonymisation, salariés de la société ayant accès aux données des clients, et mesures diverses.

Les critères de souveraineté

La grille classe quatre critères de souveraineté des données, par ordre de priorité :

1. le lieu d'hébergement des données des clients (France, voire Union européenne) ;
2. la nationalité de l'entreprise détenant les serveurs qui hébergent les données des clients (française, voire européenne), **en excluant toute entreprise dont le droit national contient des dispositions extraterritoriales permettant à ses autorités de réquisitionner des données stockées sur des serveurs appartenant aux entreprises de sa nationalité, « comme les entreprises américaines en l'état actuel de leur législation »** ;
3. le lieu d'hébergement des données du grand modèle de langage (France, voire Union européenne) ;
4. la nationalité de l'entreprise détenant les serveurs qui hébergent les données du LLM (même règle qu'au point 2).

Elle ajoute un prérequis : l'outil ne doit pas réutiliser les données pour entraîner le modèle. Tous les outils audités affirment l'appliquer, mais la grille relève qu'une telle assertion d'un modèle propriétaire n'est « en général ni prouvable ni vérifiable ». Pour écarter ces risques, elle voit deux voies : l'anonymisation totale préalable des données, ou l'installation du LLM sur les systèmes du cabinet. Le guide du CNB sur la déontologie et l'intelligence artificielle (avril 2026)² élargit la première voie à la pseudonymisation.

Ce que déclarent les quinze fiches

Nous avons lu les quinze fiches du PDF publié par le CNB et relevé deux champs. Le libellé est celui du CNB ; « N/C » signifie non communiqué, « N/A » non applicable.

Outil	Société hébergeant le LLM	Pseudonymisation / anonymisation
Doctrine	Américaine (Microsoft)	Pseudonymisation
Lefebvre Dalloz	Américaine (Microsoft)	N/C

Jimini	Française (Mistral) et américaine (Microsoft)	Pseudonymisation
Juri'Predis	Américaine (Microsoft)	Pseudonymisation
Jus Mundi	Américaine (Microsoft)	Anonymisation
Lamy Liaisons	Américaine (Microsoft)	N/C
Legora	Américaines (Microsoft, AWS et Amazon Bedrock)	N/C
Lexbase	Américaine (Microsoft) et en local chez Lexbase	Anonymisation
LexisNexis	Américaines (Microsoft, AWS et Amazon Bedrock)	Non, mais suppression au bout de quelques minutes / jours
Notae	Américaine (Google)	N/C
Ordalie	Américaine (Microsoft) et en local chez Ordalie	N/C
Septeo	Américaine (Microsoft)	N/C
Cedie	N/A (en local sur la machine de l'utilisateur)	N/A
Predictice	Américaines	N/C (peut-être précisé dans le « Trust Center »)
Haiku	Américaines	N/C

Source : CNB, *Grille de lecture*, juin 2025, v1.2, fiches de synthèse (p. 29 à 43 du PDF). Relevé et décompte par nos soins, 24 et 25 octobre 2026. Outils dans l'ordre de la grille.

- **Hébergement du LLM.** 14 fiches sur 15 citent une société américaine parmi celles qui hébergent le LLM : 11 ne citent que des sociétés américaines, 3 en citent une à côté d'un modèle français ou d'une installation en local (Jimini, Lexbase, Ordalie). La quinzième, Cedie, fonctionne sur la machine de l'utilisateur.
- **Pseudonymisation ou anonymisation.** 5 fiches déclarent un mécanisme (3 pseudonymisation, 2 anonymisation), 1 répond « non » en mettant en avant une suppression rapide (LexisNexis), 8 indiquent « N/C » et 1 « N/A ». Au total, **9 fiches sur 15 ne déclarent rien** dans ce champ.

Ce n'est pas un classement. Les fiches reflètent ce que les éditeurs ont déclaré au groupe de travail en 2025 (la grille le précise : « sous couvert d'évolution et en fonction des éléments déclarés, trouvés et/ou transmis par les sociétés »), et plusieurs éditeurs ont publié davantage depuis. C'est la

photographie du marché qu'a prise un régulateur attentif : presque tous les outils dépendent d'un cloud américain pour le modèle, et la plupart ne disent pas si le modèle reçoit les noms en clair.

Au-delà du barreau

Les autres régulateurs en ont dit moins, et moins précisément. Les lignes directrices du CEPD sur la pseudonymisation restent un projet soumis à consultation, et nous n'avons trouvé aucune délibération de la CNDP propre à l'IA générative (section 1). Les règles de la profession placent la barre haut : le Règlement intérieur national dit du secret professionnel qu'il est « général, absolu et illimité dans le temps » (article 2.1)¹⁶, et le guide du CNB d'avril 2026 recommande de « pseudonymiser ou anonymiser systématiquement les données avant toute utilisation d'un outil d'IA générative », tout en prévenant que la pseudonymisation ne s'applique pas à toutes les tâches et qu'une personne peut être réidentifiée par recoupement². La fiche en dix champs est la liste de contrôle la plus détaillée qu'un ordre professionnel ait publiée, et elle sert bien au-delà des avocats.

3 La détection des données personnelles françaises et marocaines

Une pseudonymisation ne vaut que par la détection qui la précède : un nom que l'outil ne voit pas arrive en clair dans l'IA. La plupart des bancs d'essai publics de détection de données personnelles sont en anglais, avec des identifiants américains. Nous avons construit HT-Bench pour mesurer les identifiants des dossiers français et marocains : CIN, ICE, IF et RC, CNSS, RIB, NIR, SIREN et SIRET, IBAN, numéros de téléphone, et noms de personnes du Maghreb et de France.

Résultats par pack métier

HT-Bench 0.2.0, jeu de test tenu à l'écart : 366 documents synthétiques, 3 420 valeurs étiquetées dont 2 523 identifiants directs, et 3 029 leurres (montants, dates, articles de loi, mots courants) qui doivent rester lisibles. Connecteur Hidden Tokens 2.28.0 face à Microsoft Presidio 2.2.358 avec ses reconnaisseurs par défaut, passage du 25 octobre 2026, mêmes documents, même code de notation¹¹. La colonne 2.27.0 reprend le passage précédent, sur le même jeu de test.

Pack	Docs	Hidden Tokens 2.28.0, identifiants directs	2.27.0	Hidden Tokens, précision	Presidio, identifiants directs	Presidio, précision
telecom-ma Télécom (Maroc)	36	100,0 % (242/242)	100,0 %	82,7 %	81,8 % (198/242)	51,8 %
core Socle (tous documents)	42	100,0 % (322/322)	99,7 %	93,5 %	87,0 % (280/322)	59,9 %
banque-assurance Banque et assurance	96	100,0 % (652/652)	99,5 %	81,7 %	64,4 % (420/652)	54,1 %
expertise-comptable Expertise comptable	84	100,0 % (599/599)	99,3 %	95,5 %	81,8 % (490/599)	60,6 %
sante Santé	36	100,0 % (204/204)	96,1 %	82,7 %	86,3 % (176/204)	50,5 %
juridique-fr Juridique (France)	36	100,0 % (276/276)	96,0 %	87,2 %	82,2 % (227/276)	50,3 %
rh-paie RH et paie	36	100,0 % (228/228)	92,1 %	89,3 %	72,4 % (165/228)	51,8 %
Tous les packs	366	100,0 % (2523/2523)	98,2 %	87,4 %	77,5 % (1956/2523)	54,7 %

Rappel sur les identifiants directs ; une valeur ne compte comme détectée que si chacune de ses lettres et chacun de ses chiffres est codé. En rouge : packs 2.27.0 sous la cible de 99 % par pack que Hidden Tokens

s'est fixée. Source : [ht-bench-2.28.0.json](#) (2.27.0 : [ht-bench-2.27.0.json](#)).

En résumé : Hidden Tokens 2.28.0 détecte au moins 99 % des identifiants directs dans chaque pack sur HT-Bench (100 % sur le jeu de test 2.28.0, 2523/2523), avec une précision de 87,4 % et 5,8 % de leurres codés. Presidio par défaut détecte 77,5 % des mêmes valeurs, avec une précision de 54,7 % et 37,6 % de leurres codés. C'est un banc synthétique passé par l'éditeur de l'outil : ce n'est pas un résultat sur de vrais dossiers.

Ce que ratent les détecteurs génériques

- **Les formats locaux.** Les reconnaisseurs par défaut de Presidio n'en comptent aucun pour les formats marocains. Il a repéré 69,8 % des numéros de CIN (134/192), 66,0 % des ICE (105/159) et 5,5 % des RIB (6/109), surtout grâce à des reconnaisseurs écrits pour d'autres pays (permis de conduire américain, PAN indien, compte bancaire américain). Les mêmes reconnaisseurs signalent presque n'importe quel long nombre : c'est pourquoi Presidio a codé 37,6 % des leurres.
- **Les raisons sociales.** Presidio par défaut n'étiquette pas les organisations : 1,8 % des noms de société repérés (5/279). Dans un dossier commercial, le nom de la contrepartie suffit souvent à identifier le client.
- **Le NIR.** Presidio a repéré 32,5 % des numéros de sécurité sociale (13/40). Écrits avec des espaces ou un tiret avant la clé, ils étaient aussi le point faible de Hidden Tokens en 2.27.0 (26/40) ; la 2.28.0 les détecte tous les 40.
- **Ce que Hidden Tokens rate encore** se trouve hors des identifiants directs : adresses postales 92,4 % (170/184), dates de naissance 79,1 % (68/86) et numéros de chèque 75,0 % (36/48).

Presidio est une boîte à outils faite pour être complétée de reconnaisseurs sur mesure. Cette comparaison mesure ses réglages par défaut face à un produit fini, pas ce qu'une équipe d'ingénieurs pourrait bâtir dessus. Elle montre ce qu'obtient un cabinet qui déploie un détecteur générique sur des dossiers français ou marocains sans ce travail.

Méthode et limites

- **Données synthétiques uniquement.** Les noms sont des prénoms et noms courants du Maroc, du Maghreb et de France recombinaison au hasard, en caractères latins ; les numéros sont aléatoires, avec des clés de contrôle valides. Aucun vrai dossier n'a servi.
- **Jeu de test tenu à l'écart, avec une réserve.** Le moteur est réglé sur un jeu de développement distinct ; les chiffres publiés viennent du jeu de test (graine `ht-bench-v0-test`, SHA-256 `4881c8d2...e10b`). Les corrections de la 2.28.0 ont été faites sur le jeu de développement, mais les écarts du jeu de test en 2.27.0 avaient été publiés avant, et les deux jeux sont tirés des mêmes modèles de documents. Le jeu de test montre que les corrections tiennent sur des valeurs et des tournures sur lesquelles le moteur n'a pas été réglé ; il ne dit pas comment le moteur traite des formats pour lesquels personne n'a encore écrit de modèle.
- **Pas indépendant.** HT-Bench est le banc de Hidden Tokens, passé par son éditeur. Le code, le générateur de données et les résultats bruts sont publics pour que chacun puisse le rejouer ; il n'a pas été audité par un tiers.

- **Hors champ** : l'écriture arabe, les documents scannés et les erreurs d'OCR, les mises en page Excel et Word (couvertes par un test de fuite distinct), et le pack notariat `notariat-ma`, livré avec le connecteur mais non mesuré : HT-Bench 0.2.0 ne contient pas d'actes notariés.

Méthode complète, tableaux par type et reproduction : hiddentokens.com/fr/banc-d-essai¹².

4 Carte du marché : quatre façons de tenir les données des clients à l'écart de l'IA

Quatre catégories de produits répondent à la question « comment utiliser l'IA sur des dossiers confidentiels ? ». On les vend souvent comme des substituts ; elles se lisent mieux comme des couches. Ce qui les distingue, c'est ce que l'IA reçoit au bout du compte, et qui détient le moyen de le lire.

Catégorie	Ce que l'IA reçoit	Vraies valeurs remises dans la réponse	Qui l'exploite	Exemples (vérifiés le 25 oct. 2026)
1. Offres entreprise des fournisseurs d'IA	Les données en clair, sous contrat : pas d'entraînement par défaut, maîtrise de la conservation, parfois une région d'hébergement	Sans objet	Le fournisseur d'IA ; l'administrateur du cabinet règle les options	ChatGPT Entreprise et Business ¹⁷ , Claude Team et Enterprise ¹⁸ , Microsoft Copilot ¹⁹ , Gemini pour Workspace ²⁰ , Mistral Vibe ²¹
2. DLP et CASB pour l'IA générative	Rien si le prompt est bloqué ; le prompt amputé de ce qui a été retiré, s'il est caviardé	En général non	Les équipes informatique et sécurité, pour toute l'organisation	Microsoft Purview (avertit ou bloque) ²² , Nightfall (caviarde ou bloque) ²³ , Strac (annonce aussi une pseudonymisation réversible) ²⁴
3. API de détection et de caviardage	Ce qu'envoie la chaîne construite dessus	Possible, si on le développe	Des développeurs, dans les applications du cabinet	Presidio (open source ; opérateurs encrypt / decrypt) ²⁵ , Limina, ex-Private AI (conteneurs ; API de réidentification) ²⁶

4. Couches de pseudonymisation réversible	Des codes cohérents à la place des noms et identifiants	Oui, automatiquement	Un proxy, un service de coffre ou le poste de travail, selon le produit	LLM Guard (Vault + Deanonymize) ²⁷ , garde-fou Presidio de LiteLLM ²⁸ , Skyflow LLM Privacy Vault ²⁹ , Protecto ³⁰ , Hidden Tokens (poste de travail)
--	---	----------------------	---	---

Catégories et exemples sont les nôtres. Chaque exemple est décrit d'après la documentation publique de son éditeur, lue le 25 octobre 2026 ; la liste n'est pas exhaustive.

Offres entreprise : ce que disent les pages des éditeurs

- **OpenAI** (page mise à jour le 8 janvier 2026) : pas d'entraînement sur les données professionnelles par défaut ; les administrateurs règlent la conservation ; les conversations supprimées disparaissent sous 30 jours. La résidence des données au repos est proposée pour Enterprise, Edu, Healthcare et l'API ; ChatGPT Business n'est pas cité¹⁷.
- **Anthropic** (pages mises à jour de juin à août 2026) : pas d'entraînement sur les saisies et sorties des produits commerciaux par défaut ; Enterprise propose une conservation personnalisée, 30 jours minimum ; la zéro conservation ne vaut que pour les API éligibles et Claude Code, pas pour le chat claude.ai ; les données sont stockées aux États-Unis¹⁸.
- **Microsoft Copilot**, ex-Microsoft 365 Copilot (page mise à jour le 8 octobre 2026) : prompts et réponses ne servent pas à entraîner les modèles de fondation ; le service respecte l'EU Data Boundary, à ceci près que les modèles fournis par Anthropic en sont pour l'instant exclus¹⁹.
- **Google Gemini pour Workspace** (mis à jour le 6 octobre 2026) : pas d'entraînement sur les données du client sans autorisation préalable ; l'administrateur fixe la conservation, de 90 jours à illimitée ; le traitement peut être cantonné aux États-Unis ou à l'UE²⁰.
- **Mistral**, dont Le Chat est devenu Vibe le 12 août 2026 : les clients Enterprise sont exclus de l'entraînement par défaut ; les données sont hébergées dans l'UE par défaut ; la zéro conservation n'est pas disponible pour le chat, quelle que soit l'offre²¹.

Chaque offre de cette catégorie améliore le sort des données *après* leur réception par le fournisseur. Aucune ne change le fait qu'il les reçoit en clair. C'est pourquoi la grille du CNB pose une question distincte sur la pseudonymisation, et pourquoi les trois autres catégories existent.

Points de vigilance

- **Où est la clé.** Dans la catégorie 4, un coffre exploité par un éditeur et une clé sur le poste de travail sont deux réponses très différentes aux critères 2 et 4 du CNB.
- **La détection dans vos langues.** La documentation de Presidio indique que, dans sa configuration par défaut, il contient des reconnaisseurs et des modèles pour l'anglais²⁵ ; la section 3 montre ce que cela donne sur des dossiers français et marocains.

- **Les noms changent vite.** Rien qu'en 2026, Private AI est devenu Limina (mars), Le Chat est devenu Vibe (août), Microsoft 365 Copilot est devenu Microsoft Copilot, et la documentation de Presidio est passée à un projet communautaire²⁵²⁶. Datez chaque affirmation sur laquelle vous appuyez.

5 Recommandations : une liste de contrôle en 10 points

Pour un cabinet d'avocats, un cabinet d'expertise comptable, une direction des ressources humaines ou toute entreprise qui confie des dossiers de clients ou de salariés à l'IA. Aucun de ces points n'exige un produit particulier.

- 1 **Cartographier d'abord les usages réels.** Demander aux équipes quels outils d'IA elles utilisent, sur quels dossiers, depuis quels appareils. Un usage non cartographié est un usage qu'on ne gouverne pas.
- 2 **Lister ce qui ne doit jamais sortir en clair.** Identité des clients, numéros d'identité et de compte, données de santé, salaires, numéros de dossier. Écrire la liste : elle devient la configuration de l'outil retenu.
- 3 **Choisir une offre professionnelle, jamais un compte personnel,** pour chaque outil d'IA autorisé : entraînement sur vos données désactivé, durée de conservation connue, contrôles d'administration activés.
- 4 **Poser à chaque éditeur les dix questions du CNB,** quelle que soit votre profession : qui héberge le modèle et les données, sous quelle juridiction, avec quel chiffrement, si les données sont réutilisées, et si les noms sont pseudonymisés avant que le modèle ne les voie. « N/C » est aussi une réponse.
- 5 **Pseudonymiser en amont de l'IA quand c'est possible,** comme le recommande le guide du CNB de 2026 : le modèle reçoit des codes cohérents, la clé reste chez vous, et les vraies valeurs sont remises ensuite.
- 6 **Garder la clé soi-même.** La pseudonymisation ne protège que si celui qui reçoit les données codées ne peut pas les réidentifier. Une clé tenue par le fournisseur d'IA ou l'éditeur de l'outil protège peu.
- 7 **Tester la détection sur vos propres documents.** Prendre dix dossiers types, les faire passer dans l'outil, et chercher chaque nom et chaque numéro resté lisible. Demander à l'éditeur son rappel par format, sur vos langues et vos identifiants.
- 8 **Fermer les portes dérobées.** Texte tapé directement dans une conversation, fichiers glissés dans un navigateur, dossiers synchronisés, extensions et agents ayant accès aux fichiers. Décider quelle voie est autorisée, bloquer ou surveiller les autres.
- 9 **L'écrire.** Une charte IA, un registre des traitements, une analyse d'impact quand le RGPD ou la loi 09-08 l'exige, et l'information du client que demandent vos règles professionnelles.
- 10 **Garder la preuve.** Pouvoir montrer à un client, à un DPO ou à un ordre ce que l'IA a réellement reçu, dossier par dossier, et le revoir au moins une fois par an.

6 La place de Hidden Tokens (déclaration d'intérêt)

Ce rapport est publié par Tardigrade SARL, qui commercialise Hidden Tokens. Hidden Tokens appartient à la quatrième catégorie ci-dessus : un logiciel de bureau qui code les noms, identifiants et montants sensibles sur l'ordinateur de l'utilisateur avant que Claude, ChatGPT, Gemini ou Copilot ne les lisent, garde la clé sur cet ordinateur ou sur le partage de fichiers de l'organisation, et remet les vraies valeurs dans le Word ou l'Excel produit par l'IA. La section 3 utilise son banc d'essai : lisez ces chiffres comme la mesure d'un éditeur.

Ce qu'il ne fait pas : il ne protège que ce qui passe par lui (le texte tapé dans une conversation sans son extension n'est pas couvert), il signale les identifiants indirects sans tous les réécrire, et il outille la conformité sans la certifier. Ses réponses aux dix champs du CNB, avec leurs limites, sont dans le dossier de conformité CNB.

7 Méthode, conflits d'intérêts, comment citer

- **Sources.** Sources primaires uniquement : textes, publications des régulateurs, décisions de justice et documentation des éditeurs. Chaque fait sur un produit tiers est daté ; les produits changent, et nous corrigerons cette page si on nous le signale à info@hiddentokens.com.
- **Les chiffres CNB** sont notre propre décompte à partir du PDF publié par le CNB, pas des chiffres publiés par le CNB.
- **Les chiffres de détection** viennent du banc d'essai de Hidden Tokens (voir ses limites en section 3).
- **Pas un avis juridique.** Ce rapport décrit le droit et le marché ; il ne remplace pas un conseil sur la situation d'un cabinet.

Comment citer :

Hidden Tokens, *État de la confidentialité de l'IA 2026*, 1^{re} édition, 25 octobre 2026, <https://hiddentokens.com/fr/etat-confidentialite-ia-2026/>

La version anglaise, State of AI confidentiality 2026, dit la même chose.

Sources

1. CNB, groupe de travail Intelligence artificielle, *Grille de lecture — synthèse des consultations effectuées par le groupe de travail sur l'intelligence artificielle*, 1^{re} édition, juin 2025, v1.2. cnb.avocat.fr/medias/cnb-grille-lecture-intelligence-artificielle-2025-68f775b81ba5b2.65545272.pdf
2. CNB, commission Règles et usages, *La déontologie et l'intelligence artificielle*, 1^{re} édition, avril 2026. cnb.avocat.fr/medias/guide-deontologie-ia-6a4cef5ea1d577.28126431.pdf
3. Eurostat, « 20% of EU enterprises use AI technologies », 11 décembre 2025 ; et *Use of artificial intelligence in enterprises*, Statistics Explained (données 2025). ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20251211-2
4. Insee, *Insee Première* n° 2120, juillet 2026. www.insee.fr/fr/statistiques/9025878

5. Règlement (UE) 2016/679 (RGPD), articles 4, 25, 28, 32, chapitre V, considérant 26. eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679
6. Comité européen de la protection des données (CEPD, EDPB), lignes directrices 01/2025 sur la pseudonymisation, adoptées le 16 janvier 2025, version soumise à consultation. www.edpb.europa.eu/documents/guideline/guidelines-012025-on-pseudonymisation_en
7. CJUE, 4 septembre 2025, Contrôleur européen de la protection des données c. Conseil de résolution unique, C-413/23 P (communiqué n° 107/25). curia.europa.eu/site/upload/docs/application/pdf/2025-09/cp250107en.pdf
8. Loi n° 09-08 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel (dahir n° 1-09-15 du 18 février 2009, BO n° 5714 du 5 mars 2009), et décret n° 2-09-165, portail Adala du ministère de la Justice. adala.justice.gov.ma/api/uploads/2024/04/30/Protection%20des%20personnes%20physiques-1714464099884.pdf
9. Loi n° 71-1130 du 31 décembre 1971 portant réforme de certaines professions judiciaires et juridiques, article 66-5. www.legifrance.gouv.fr/loda/article_lc/LEGIARTI000006512415
10. Code pénal, article 226-13. www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000006417945
11. Hidden Tokens, résultats bruts HT-Bench 0.2.0, jeu de test, connecteur 2.28.0, 25 octobre 2026 (JSON) ; passage précédent, connecteur 2.27.0 : hiddentokens.com/benchmark/ht-bench-2.27.0.json ; hiddentokens.com/benchmark/ht-bench-2.28.0.json
12. Hidden Tokens, *Banc d'essai : détection par pack métier, et test de fuite*. hiddentokens.com/fr/banc-d-essai/
13. Règlement (UE) 2024/1689 sur l'intelligence artificielle, article 113. eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32024R1689
14. Règlement (UE) 2026/1744 du 8 juillet 2026 (Omnibus numérique sur l'IA), JO L du 24 juillet 2026. eur-lex.europa.eu/eli/reg/2026/1744/oj/eng
15. Commission européenne, COM(2025) 837, procédure 2025/0360(COD), en cours. eur-lex.europa.eu/procedure/EN/2025_360
16. CNB, Règlement intérieur national de la profession d'avocat, version consolidée au 7 juillet 2026, article 2.1. cnb.avocat.fr/medias/file/cnbrinmaj-08072026-6a4e79014f6b41.49207715.pdf
17. OpenAI, *Enterprise privacy* (mise à jour le 8 janvier 2026) et *Business data*, consultés le 25 octobre 2026. openai.com/enterprise-privacy/
18. Anthropic Privacy Center : entraînement (mis à jour le 18 août 2026), conservation, zéro conservation, emplacement des serveurs (15 juin 2026), consultés le 25 octobre 2026. privacy.claude.com/en/articles/7996868-is-my-data-used-for-model-training
19. Microsoft Learn, *Data, privacy, and security for Microsoft Copilot* (mis à jour le 8 octobre 2026). learn.microsoft.com/en-us/copilot/microsoft-365/microsoft-365-copilot-privacy
20. Google, *Generative AI in Google Workspace Privacy Hub* (mis à jour le 6 octobre 2026). knowledge.workspace.google.com/admin/generative-ai/generative-ai-in-google-workspace-privacy-hub
21. Mistral AI, centre d'aide : « Le Chat is now Vibe » (12 août 2026), entraînement (3 septembre 2026), zéro conservation et stockage des données (12 août 2026). help.mistral.ai/en/articles/682992-le-chat-is-now-vibe
22. Microsoft Learn, *Microsoft Purview data security and compliance protections for generative AI apps* (27 mai 2026). learn.microsoft.com/en-us/purview/ai-microsoft-purview
23. Nightfall AI, pages solutions « Securely use GenAI » et « Prevent data leakage to shadow AI », consultées le 25 octobre 2026. www.nightfall.ai/solutions/securely-use-genai
24. Strac, billet « ChatGPT DLP » (1^{er} octobre 2026). www.strac.io/blog/chatgpt-dlp-data-loss-prevention
25. Presidio, documentation (langues, anonymiseur) ; presidio-analyzer 2.2.364 sur PyPI (22 juillet 2026) ; version mesurée en section 3 : 2.2.358. presidio.dataprivacystack.org/
26. Limina (ex-Private AI), site et documentation (getlimina.ai, docs.getlimina.ai), communiqué de changement de nom du 5 mars 2026. www.getlimina.ai/en
27. Protect AI, LLM Guard, scanners Anonymize et Deanonymize. protectai.github.io/llm-guard/input_scanners/anonymize/
28. LiteLLM, garde-fou Presidio de masquage de PII. docs.litellm.ai/docs/proxy/guardrails/pii_masking_v2

29. Skyflow, documentation *LLM privacy*. docs.skyflow.com/docs/fundamentals/patterns/llm-privacy.

30. Protecto, page produit *Privacy Vault*. www.protecto.ai/product/privacy-vault/

Ce rapport décrit le droit et le marché ; il ne vaut pas avis juridique. © 2026 Tardigrade SARL. Claude est une marque d'Anthropic ; ChatGPT, Gemini, Copilot, Le Chat et les autres produits cités sont des marques de leurs détenteurs. Hidden Tokens n'est affilié à aucun d'eux, ni au Conseil national des barreaux.