LE POINT SUR...

Confidentialité des données et usage de l'IA

Où transitent et où sont stockées les données de l'entreprise

De nombreuses entreprises considèrent la question réglée dès lors que l'option autorisant le fournisseur à utiliser leurs contenus pour améliorer son modèle a été désactivée. La mesure est nécessaire, mais elle ne répond qu'à une question sur cinq. L'essentiel porte sur le trajet des données : par où elles transitent, où elles sont effectivement traitées, et combien de temps elles sont conservées. Ces réponses conditionnent ce qui peut être confié à un outil d'IA, et ce qui ne le peut pas.

Par où passent les données ?

Le texte saisi dans un outil d'IA ne reste pas sur le poste de travail. Il est transmis à un fournisseur, traverse une chaîne technique, puis demeure conservé un certain temps. Cette chaîne comporte trois étages souvent confondus, qui peuvent se situer dans trois pays différents.

- La plateforme. L'interface de travail, qui conserve les historiques et les documents qui lui sont confiés.
- Le modèle. Le moteur qui produit la réponse, souvent fourni par un tiers, distinct de l'éditeur de la plateforme.
- Le calcul. L'endroit physique où le document est effectivement traité, qui n'est pas nécessairement celui où la plateforme est hébergée.

Un éditeur français peut ainsi héberger son interface en France, s'appuyer sur un modèle américain et faire réaliser les calculs ailleurs. La situation n'a rien d'anormal, mais sans ces trois réponses, l'entreprise ne peut pas documenter correctement le niveau de risque ni déterminer quels usages peuvent être autorisés.

Toutes les offres ne se valent pas

Type d'offre	Garanties habituelles	Usages envisageables
Compte gratuit ou personnel	Aucun engagement contractuel ; contenus parfois utilisés pour l'entraînement par défaut	Aucun usage professionnel
Offre professionnelle ou entreprise	Absence d'entraînement, contrat de traitement des données, conservation limitée, gestion centralisée des comptes	Documents internes, données clients peu sensibles
Accès via un hébergeur choisi	Choix de la région de calcul, options sans conservation des échanges	Données clients sous confidentialité, données RH
Solution installée en interne ou hébergée auprès d'un fournisseur répondant à des exigences de sécurité et de souveraineté adaptées	cette architecture peut réduire certains risques, sous réserve de vérifier la localisation, les accès administrateurs, les sous-traitants, les journaux, les sauvegardes, la télémétrie et les conditions contractuelles	Données sensibles, secret professionnel, santé

Un même éditeur propose fréquemment plusieurs de ces niveaux sous un nom commercial identique : c'est l'offre souscrite qui détermine les garanties, non la marque.

© ATH. Tous droits réservés.

Les cinq questions à poser au fournisseur

Point à vérifier	Question à poser
L'entraînement	Nos contenus servent-ils à améliorer votre modèle, et le réglage est-il désactivé par défaut pour notre compte ?
La plateforme	Où sont hébergés l'interface, nos historiques et les documents qui y sont déposés ?
Le modèle	Quel modèle est utilisé, fourni par quelle société, et sous quel contrat ?
Le lieu de calcul	Dans quel pays nos documents sont-ils traités, et cet engagement figure-t-il par écrit ?
La conservation	Combien de temps nos échanges sont-ils conservés, et qui peut y accéder ?

La conservation est souvent insuffisamment documentée. Même lorsque les contenus ne servent pas à entraîner le modèle, le fournisseur peut conserver temporairement certains échanges ou journaux pour des motifs de sécurité, de support ou de détection des abus. Les conditions d'un éventuel accès humain doivent être vérifiées dans la documentation contractuelle et technique du service.

Une conservation ou un accès non maîtrisé peut rendre l'outil inadapté à certaines données couvertes par une obligation renforcée de confidentialité.

Et si le fournisseur est américain ?

Une société américaine demeure soumise au droit américain, qui peut l'obliger à communiquer des données qu'elle contrôle, y compris lorsque celles-ci sont stockées en Europe.

Le stockage en Europe n'est pas neutre pour autant : l'accès à ces données y est protégé par le droit européen, ce qui rend la démarche nettement plus contraignante que pour des données situées sur le sol américain. La localisation reste donc un critère utile, sans constituer une barrière absolue.

Lorsqu'une organisation américaine est valablement certifiée au titre du *Data Privacy Framework* et que la certification couvre le traitement concerné, ce mécanisme peut servir de base au transfert au titre du chapitre V du RGPD. Il ne dispense toutefois pas l'entreprise de respecter les autres obligations du RGPD : base légale, information, sécurité, minimisation, encadrement du prestataire et documentation du traitement.

Si le destinataire ne bénéficie pas d'une décision d'adéquation applicable, le transfert doit généralement reposer sur un mécanisme approprié, comme les clauses contractuelles types, complété par une analyse du contexte du transfert et, si nécessaire, par des mesures supplémentaires de protection.

Il existe par ailleurs des offres dites souveraines et une qualification SecNumCloud délivrée par l'ANSSI.

Le recours à un fournisseur américain n'est pas à exclure par principe : il s'agit d'un critère à pondérer selon la sensibilité des données.

* * * * * *

Concrètement pour l'entreprise

Avant le choix d'un outil : classer les données

Exemples d'usages	*Règle applicable*
Niveau libre	
Textes publics, brouillons sans nom, recherches générales, supports de communication	Tout outil autorisé par l'entreprise
Niveau sous contrat	
Données clients, devis, comptes rendus internes, données RH courantes	Uniquement l'outil professionnel validé
Niveau jamais	
Données de santé, secrets de fabrication, dossiers sous secret professionnel ou accord de confidentialité, mots de passe	Aucun outil externe ; solution interne uniquement

© ATH. Tous droits réservés.

Auprès du fournisseur

- **Exiger les cinq réponses par écrit**, dans le contrat ou son annexe, et non dans un argumentaire commercial.
- **Demander les documents utiles :** contrat de traitement des données, liste des sous-traitants, politique de conservation et certifications - ISO 27001 en général, HDS pour les données de santé, SecNumCloud en cas d'exigence de souveraineté.
- **Obtenir l'engagement d'être informé** de tout changement de fournisseur de modèle ou de zone d'hébergement : sans cette clause, les réponses obtenues ne valent que pour aujourd'hui.

En interne

- **Interdire les comptes personnels gratuits** pour tout ce qui touche à l'activité : aucun engagement contractuel n'existe, alors que la responsabilité demeure celle de l'entreprise.
- **Formaliser une charte IA.** Elle énonce les outils autorisés, les données admises pour chacun, et la conduite à tenir en cas d'erreur de manipulation.
- **Pseudonymiser avant transmission.** Remplacer les noms par des initiales ou des codes constitue un geste simple qui réduit fortement le risque.
- **Analyser chaque usage d'un outil d'IA et mettre à jour le registre des traitements lorsque l'usage constitue un traitement de données personnelles devant y être documenté :** finalité, données concernées, hébergement, durée de conservation.
- **Pour les données les plus sensibles**, étudier une solution hébergée en interne ou dans un cloud qualifié, voie recommandée par la CNIL.

La charte IA doit préciser les outils autorisés, les données pouvant être utilisées, les contrôles à réaliser et la conduite à tenir en cas d'erreur. Lorsqu'elle contient des règles générales et permanentes applicables aux salariés ou prévoit des sanctions disciplinaires, elle doit être intégrée au règlement intérieur ou prendre la forme d'une note de service soumise aux mêmes formalités.

Dans les entreprises qui ne sont pas soumises à l'obligation d'établir un règlement intérieur, la charte peut être diffusée comme une règle interne et servir de support aux consignes données aux salariés. Elle ne dispense pas l'employeur de respecter les règles disciplinaires applicables. Les obligations imposées aux prestataires doivent être prévues séparément dans leurs contrats.

Les points de vigilance

- La saisie de données personnelles dans un outil d'IA constitue un traitement de données. L'entreprise doit en déterminer la finalité, la base légale et les mesures de protection. Le fournisseur peut également avoir des obligations propres, selon son rôle dans le traitement : sous-traitant, responsable de traitement ou autre acteur de la chaîne.
- Pseudonymiser n'est pas anonymiser : un dossier sans nom mais comportant une date de naissance, une commune et un poste reste souvent identifiable.
- Un outil validé il y a un an a pu changer de modèle ou de sous-traitant depuis : la vérification n'est pas acquise une fois pour toutes.

À savoir

Aucun outil n'est conforme en soi. La conformité s'apprécie pour un usage donné, dans un contexte donné : le même service peut convenir à des comptes rendus internes et se révéler inadapté à un dossier client sous accord de confidentialité.

Les réponses obtenues d'un fournisseur méritent d'être datées et conservées. Elles constituent, en cas de question d'un client ou de contrôle, la trace de la diligence accomplie.

À retenir

1. Trois étages sont à distinguer : la plateforme de travail, le modèle qui répond, et le lieu où le calcul est réalisé.
2. La désactivation de l'entraînement ne règle qu'une question sur cinq ; la conservation temporaire des échanges est souvent insuffisamment documentée.
3. La marque ne détermine pas les garanties : seules comptent l'offre souscrite et le contrat.
4. Classer les données en trois niveaux avant de choisir un outil rend rapides toutes les décisions suivantes.
5. Une charte IA doit être diffusée et articulée avec les règles internes applicables. Lorsqu'elle contient des obligations générales et permanentes ou prévoit des sanctions disciplinaires, elle doit respecter les formalités applicables au règlement intérieur ou à la note de service.

Pour aller plus loin

CNIL : questions-réponses sur l'utilisation d'un système d'IA générative.

ANSSI : recommandations de sécurité pour un système d'IA générative.

Liste officielle des sociétés certifiées Data Privacy Framework : dataprivacyframework.gov/list

© ATH. Tous droits réservés.