

Le CEPD met en lumière l'anonymisation et le moissonnage pour l'IA générative et adopte la version finale des lignes directrices sur la chaîne de blocs

9.7.2026 - | Commission Nationale de l'Informatique et des Libertés

Le 7 juillet 2026, le CEPD a adopté des lignes directrices sur l'anonymisation et des lignes directrices sur le moissonnage dans le contexte de l'IA générative. Il a aussi adopté la version finale de ses lignes directrices sur le traitement des données au moyen des technologies de la chaîne de blocs.

Cet article a été traduit automatiquement à partir [du communiqué du CEPD](#).

Comprendre les données anonymes

Les nouvelles lignes directrices **du CEPD clarifient la notion de données anonymes**, en tenant également compte de l'arrêt de la Cour de justice de l'Union européenne dans l'affaire [C-413/23 P, CEPD/CRU, du 4 septembre 2025, et d'autres arrêts de la CJUE](#).

Les données sont anonymes si elles ne concernent pas une personne physique identifiée ou identifiable. La question de savoir si tel est le cas peut varier d'une entité à l'autre.

Les informations peuvent **concerner** une personne en raison de leur **contenu, de leur finalité ou de leurs effets**. L'existence d'un tel lien peut ne pas être immédiatement évidente et pourrait nécessiter une analyse plus approfondie.

Une personne est considérée comme « **identifiée ou identifiable** » si elle **peut être distinguée des autres dans un contexte spécifique** en utilisant des moyens raisonnablement susceptibles d'être utilisés d'une manière qui permet de les traiter différemment. La question de savoir si **les moyens sont raisonnablement susceptibles d'être utilisés** dépendra du point de vue de l'entité concernée et devrait être appréciée à la lumière de tous les facteurs objectifs.

Les lignes directrices fournissent également un cadre pratique permettant aux organisations de **déterminer si l'anonymisation est un succès**. Le cadre peut être appliqué de deux manières : soit **en évaluant les différences de capacités** entre ceux qui pourraient identifier l'individu (« **approche contextuelle** »), soit, par souci de simplicité, en ne tenant pas compte **de ces différences** (« **approche simplifiée** »), si un responsable du traitement choisit de le faire. L'approche contextuelle reflète toutes les nuances de la norme juridique relative à l'anonymisation. L'approche simplifiée peut aller au-delà de la norme juridique et peut amener un responsable du traitement anonymisant à traiter les données comme si elles n'étaient pas anonymes, même si ce serait effectivement le cas pour certaines entités concernées, mais cette approche peut être plus pratique et fournir une plus grande confiance dans le fait que les données sont réellement anonymes.

Le cadre utilise **3 critères** pour tester si les données sont anonymes :

1. **pas d'individualisation des enregistrements,**

2. **pas de corrélation**, et
3. **pas d'inférence**.

Si les trois critères sont remplis, les données peuvent être considérées comme anonymes en toute sécurité. Si l'un de ces critères n'est pas satisfait, une analyse plus approfondie devrait être effectuée pour déterminer si les données peuvent être considérées comme anonymes.

Les lignes directrices feront l'objet d'une [consultation publique](#) jusqu'au **30 octobre 2026**, ce qui donnera aux parties prenantes la possibilité de formuler des observations et de fournir un retour d'information.

Clarifier les implications du moissonnage sur la protection des données pour le développement de l'IA

Le moissonnage (*web scraping* en anglais) est un processus automatisé d'extraction de données à grande échelle qui fonctionne souvent sans que les individus en soient conscients et qui peut présenter des risques importants pour la protection de leurs données personnelles. Dans ses **lignes directrices sur le moissonnage dans le contexte de l'IA générative**, le comité clarifie divers aspects de la conformité du moissonnage au RGPD, y compris **la base juridique de ces activités** et les **conditions dans lesquelles des catégories particulières de données peuvent être traitées dans ce contexte**.

L'IA générative est une technologie visant à créer de nouveaux contenus en apprenant des modèles à partir de données existantes. Il utilise des modèles d'apprentissage automatique spécialisés conçus pour produire une grande variété de sorties telles que du texte, de l'image ou de l'audio.

Le RGPD s'applique au moissonnage lorsqu'il inclut des opérations de traitement de données à caractère personnel, telles que **la collecte, le stockage, l'organisation et l'extraction**.

Lorsqu'on s'appuie sur le moissonnage, il convient d'accorder une attention particulière au **principe de limitation de la finalité** et au **principe de transparence**. Toutefois, en fonction de la conception précise du traitement des données, le responsable du traitement pourrait ne pas avoir à informer personnellement les personnes si cela s'avère impossible ou nécessite des efforts excessifs.

Le CEPD recommande d'extraire les données uniquement à partir de sources fiables, d'enregistrer l'horodatage et de valider les données avant de les utiliser dans le cadre d'une formation à l'IA afin de garantir le respect du **principe de précision**. Les lignes directrices donnent également des conseils sur les mesures que le responsable du traitement devrait mettre en œuvre pour se conformer au **principe de minimisation des données**.

Sur la base de l'avis [du comité européen de la protection des données sur les modèles d'IA](#), les lignes directrices fournissent des précisions et des exemples supplémentaires sur l'utilisation de la base juridique relative à l'intérêt légitime dans le contexte spécifique du moissonnage pour la formation à l'IA.

Enfin, le CEPD rappelle que **le traitement de catégories particulières de données à caractère personnel est en principe interdit**. Si le moissonnage implique de telles données, une base légale en vertu de l'article 6 du RGPD et une exception en vertu de l'article 9, paragraphe 2, du RGPD sont requises. Le CEPD suggère que l'arrêt [de la Cour dans l'affaire GC & e.a. \(C-136/17\)](#) peut être

pertinent pour **la collecte accessoire ou résiduelle** de catégories particulières de données à caractère personnel, à condition que le responsable du traitement agisse dans le « cadre de ses **responsabilités, pouvoirs et capacités** » et **mette en œuvre des mesures techniques et organisationnelles appropriées** pour empêcher la collecte et la diffusion de ces données. Le CEPD souligne qu'il n'existe pas d'exemption générale aux exigences de l'article 9 du RGPD et que chaque affaire doit être appréciée individuellement pour déterminer si le raisonnement de la Cour s'applique.

Les lignes directrices feront l'objet d'une [consultation publique](#) jusqu'au **30 octobre 2026**, ce qui donnera aux parties prenantes la possibilité de formuler des observations et de fournir un retour d'information.

Texte reference

Lignes directrices sur les chaînes de blocs finalisées après consultation publique

À la suite d'une consultation publique, le CEPD a adopté la version finale de ses lignes directrices sur les technologies des chaînes de blocs. Les lignes directrices aident les organisations qui utilisent les technologies blockchain à se conformer au RGPD. Le CEPD explique le fonctionnement des chaînes de blocs, en évaluant les différentes architectures possibles et leurs implications pour le traitement des données à caractère personnel.

Conformément à l'objectif [de la déclaration d'Helsinki](#) de renforcer le dialogue avec les parties prenantes, le comité a également publié un [rapport sur les résultats de la consultation publique spécifique](#), ainsi qu'une [version des lignes directrices intitulée « Track changes »](#).

<https://www.cnil.fr/fr/cepd-ia-generative-chaines-blocs>