#HASHTAG
CABINET D'AV0CATS

Les fiches pratiques IA

Synthèse source CNIL

Arnaud Touati

#HASHTAG
CABINET D'AV0CATS

Matthieu Chassagne

pinacle

“ Bienvenue dans notre guide pratique sur les aspects juridiques de l'intelligence artificielle.

Que vous soyez développeur, chef de projet, juriste ou décideur, ce document a été conçu pour vous accompagner dans la mise en conformité de vos projets d'IA.

À travers les 12 fiches thématiques de la CNIL, nous abordons les points clés de la réglementation, de la conception à la mise en œuvre de vos systèmes d'IA. Chaque fiche offre des conseils pratiques et des éclairages juridiques pour naviguer sereinement dans cet environnement complexe et en constante évolution.

Notre objectif : vous permettre d'innover de manière responsable, en respectant les droits des individus et les exigences légales.

Bonne lecture et bon développement ! ”

Synthèse source CNIL :
Les fiches pratiques IA

VOIR LA SOURCE

Les fiches pratiques IA

Sommaire

1 - Déterminer le régime juridique applicable

2 - Définir une finalité

3 - Déterminer la qualification juridique des acteurs

4 - Assurer que le traitement est licite

- **4.1** - Définir une base légale
- **4.2** - En cas de réutilisation des données, effectuer les tests et vérifications nécessaires

5 - Réaliser une analyse d'impact si nécessaire

6 - Tenir compte de la protection des données dans la conception du système

7 - Tenir compte de la protection des données dans la collecte et la gestion des données

8 - Mobiliser la base légale de l'intérêt légitime pour développer un système d'IA

- **8.1** - Focus sur la diffusion des modèles en source ouverte (open source)
- **8.2** - Focus sur les mesures à prendre en cas de collecte des données par moissonnage (web scraping)

9 - Informer les personnes concernées

10 - Respecter et faciliter l'exercice des droits des personnes concernées

11 - Annoter les données

12 - Garantir la sécurité du développement d'un système d'IA

INTRODUCTION

LES FICHES PRATIQUES IA

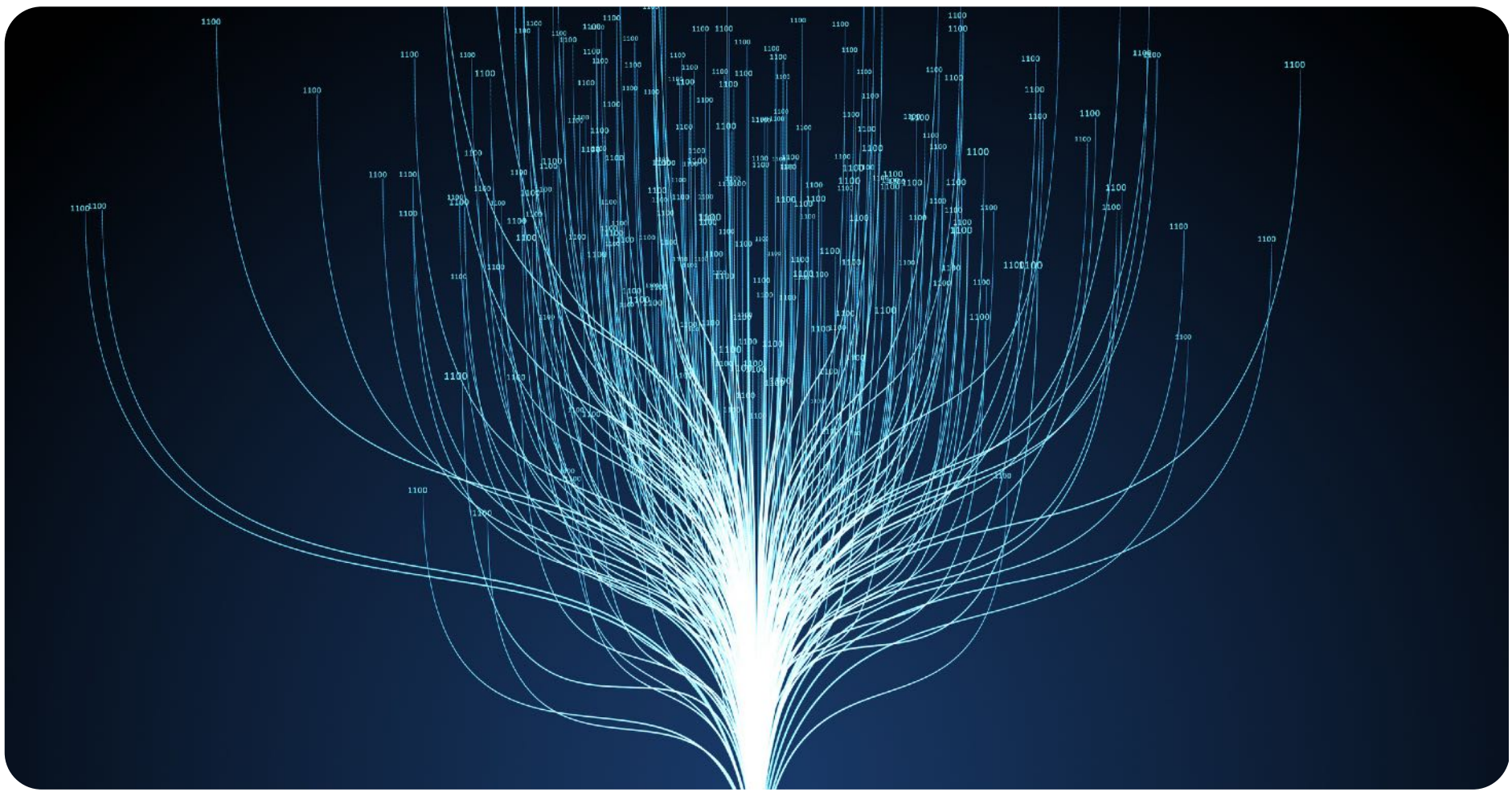
SYNTHÈSE SOURCE CNIL

L'intelligence artificielle (IA) transforme rapidement notre société, apportant des innovations majeures dans de nombreux domaines. Cependant, son essor soulève des questions cruciales en matière de protection des données, d'éthique et de sécurité.

Face à ces défis, les régulateurs ont commencé à élaborer des cadres juridiques spécifiques. En Europe, le RGPD et le projet de règlement sur l'IA (AI Act) posent les bases d'une approche réglementaire centrée sur la protection des individus et la gestion des risques.

Dans ce contexte, il est essentiel pour les acteurs de l'IA de comprendre et d'intégrer les exigences légales et éthiques dès la conception de leurs systèmes. Ce guide offre une feuille de route détaillée pour naviguer dans les aspects juridiques de l'IA, couvrant tous les aspects, de la définition du régime applicable à la sécurisation du développement.

L'objectif est de fournir un outil pratique permettant de développer et de déployer des solutions d'IA de manière responsable et conforme, dans le respect des droits fondamentaux des individus.



LES FICHES PRATIQUES IA

DÉTERMINER LE RÉGIME JURIDIQUE APPLICABLE

I. Principe général

Le régime juridique des traitements de données personnelles en lien avec l'intelligence artificielle (IA) est principalement régi par le Règlement Général sur la Protection des Données (RGPD). Ce règlement s'applique à tous les traitements de données personnelles, tant dans le secteur public que privé, à l'exception de deux régimes spécifiques :

Secteur "Police-Justice" :

- Réglementé par le titre III de la loi "Informatique et Libertés".

Défense nationale et sûreté de l'État :

- Réglementé par des dispositions spécifiques de la loi "Informatique et Libertés".

II. Champ d'application

La création et l'apprentissage des systèmes d'IA impliquant des données personnelles doivent respecter la réglementation sur la protection des données. Les phases de développement et de déploiement des systèmes d’IA constituent des traitements distincts, soumis à des régimes juridiques différents selon les cas.

III. Cas pratiques de régimes juridiques

Cas n°1 : Usage opérationnel défini dès la phase de développement

Critères :

- Usage opérationnel du système d'IA identifié dès la phase de développement.
- Finalité des traitements en phase de développement identique à celle de la phase de déploiement.
- Utilisation à des fins de prévention et détection des infractions pénales, enquêtes, poursuites, et exécution de sanctions pénales.
- Responsable du traitement en phase de développement est une "autorité compétente".

Régime juridique :

- Généralement soumis au même régime juridique (ex. : RGPD) pour les phases de développement et de déploiement.
- Peut relever du régime "Police-Justice" si les conditions sont remplies.

Cas n°2 : Usage opérationnel non défini dès la phase de développement (IA à usage général)

Exemple :

- Développement d'un modèle de reconnaissance vocale à usages multiples (identification de locuteurs, applications de traduction, etc.).

Régime juridique :

- Constitution de la base de données pour l'apprentissage du modèle relève du RGPD.
- Les traitements en phase de développement et de déploiement peuvent être soumis à des régimes différents en fonction de l'usage opérationnel final du système d'IA.

IV. Exemple de traitements différents selon les usages

1. Usage commercial (ex. : Centre commercial)

Régime juridique :

- Traité sous le RGPD pour les phases de développement et de déploiement.

2. Usage par une autorité de police

Régime juridique :

- Phase de développement sous le RGPD.
- Phase de déploiement sous le régime "Police-Justice" pour des finalités de poursuites judiciaires.

V. Conclusion

La détermination du régime juridique applicable aux traitements de données dans le cadre du développement et du déploiement de systèmes d’IA dépend de la clarté de l'usage opérationnel dès la phase de développement et de la nature de l'entité responsable du traitement. Les systèmes d’IA à usage général peuvent nécessiter une analyse spécifique pour chaque cas d'usage.

Cette synthèse présente les principaux points du document en détaillant les régimes juridiques applicables aux différents cas de traitement de données dans le contexte du développement et du déploiement des systèmes d’IA.

2

LES FICHES PRATIQUES IA

DÉFINIR UNE FINALITÉ

I. Principes de base

1. Principe de transparence :
   - Les personnes concernées doivent être informées de l'objectif du traitement des données personnelles afin de comprendre pourquoi leurs données sont collectées et comment elles seront utilisées.

2. Principe de minimisation :
   - Les données collectées doivent être adéquates, pertinentes et limitées à ce qui est nécessaire pour atteindre les objectifs du traitement.

3. Principe de limitation des durées de conservation :
   - Les données doivent être conservées pendant une durée limitée, définie en fonction de l'objectif pour lequel elles ont été collectées.

II. Définition de la finalité du traitement

La finalité du traitement correspond à l'objectif poursuivi par l'utilisation des données personnelles. Cette finalité doit être :

- Déterminée : Établie dès la définition du projet.
- Explicite : Connue et compréhensible par les personnes concernées.
- Légitime : Compatible avec les missions de l'organisme responsable du traitement.

III. Détermination de la finalité selon les scénarios

1. Usage opérationnel identifié dès la phase de développement

- Lorsque l'usage opérationnel du système d'IA est clair dès le développement, les finalités de traitement en phase de développement et de déploiement sont directement liées. Par conséquent, si la finalité en phase de déploiement est déterminée, explicite et légitime, celle en phase de développement le sera également.

Exemple : Un organisme crée une base de données de photos de rames de trains pour mesurer l'affluence en vue de déployer ce système dans les gares. La finalité est donc déterminée et explicite dès la phase de développement.

2. Usage opérationnel non défini dès la phase de développement (IA à usage général)

- Dans ce cas, il peut être difficile de définir une finalité précise. Cependant, la finalité doit être suffisamment précise et explicite en se référant :

1. Au type de système développé.
2. Aux fonctionnalités et capacités techniquement envisageables.

Exemple : Développement d'un modèle de classification d'images ou d'un modèle de langage avec des usages variés et non définis à l'avance.

Iv. Finalité du développement d'un système d'IA à des fins de recherche scientifique

Recherche scientifique : La finalité peut être moins précise en raison de la nature exploratoire de la recherche. Les traitements de données à des fins de recherche scientifique peuvent bénéficier de dérogations aux obligations habituelles (droits des personnes, durées de conservation, etc.).

LES FICHES PRATIQUES IA

DÉFINIR UNE FINALITÉ

Critères pour définir la finalité de recherche scientifique :

- Nouveauté : Le traitement doit viser à obtenir des résultats nouveaux.
- Créativité : Basé sur des hypothèses originales et non évidentes.
- Incertitude : Caractère incertain quant au résultat final.
- Systématicité : S'inscrire dans une planification et méthodologie scientifique.
- Transférabilité/Reproductibilité : Résultats pouvant être reproduits ou transférés à d'autres contextes.

Exemple : Développement d'un système d'IA pour démontrer la robustesse d'un apprentissage automatique nécessitant moins de données.

V. Bonnes pratiques

- La finalité doit inclure les capacités prévisibles les plus à risque et les fonctionnalités exclues par conception.
- Spécifier les conditions d'utilisation du système d’IA dès que possible, notamment les cas d'usage connus ou raisonnablement envisageables.
- Pour les systèmes d’IA à usage général, la transparence est essentielle et les utilisateurs doivent être informés des risques et obligations associés.

Conclusion

La définition de la finalité du traitement des données personnelles dans le développement des systèmes d’IA est cruciale pour assurer la conformité au RGPD. Elle doit être déterminée, explicite et légitime, et tenir compte des spécificités des différents types de systèmes d’IA, qu'ils soient à usage spécifique ou général, ou destinés à des fins de recherche scientifique.

3

LES FICHES PRATIQUES IA

DÉTERMINER LA QUALIFICATION JURIDIQUE DES ACTEURS

I. Introduction

Le document traite de la détermination de la qualification juridique des acteurs impliqués dans le développement et l'utilisation des systèmes d'intelligence artificielle (IA), notamment en ce qui concerne la protection des données personnelles. Il s'agit de définir les rôles et responsabilités des différents acteurs selon le (RGPD).

II. Principaux acteurs

1. Fournisseur de système d’IA :
   - Développe ou fait développer un système et le met sur le marché ou en service sous son propre nom ou marque, à titre onéreux ou gratuit.

2. Importateurs, distributeurs et utilisateurs de ces systèmes :
   - Les personnes déployant les systèmes d’IA.

3. Organismes constituant des bases de données d’apprentissage :
   - Peuvent être responsables de traitement, responsables conjoints ou sous-traitants.

III. Qualification des acteurs

La qualification des acteurs impliqués doit être déterminée au cas par cas, en fonction des objectifs et moyens du traitement de données.

3

LES FICHES PRATIQUES IA

DÉTERMINER LA QUALIFICATION JURIDIQUE DES ACTEURS

1. Responsable du traitement :

Détermine les objectifs et les moyens du traitement des données personnelles (le "pourquoi" et le "comment").

- Indices pour l'analyse :
  - Type de données collectées et utilisées.
  - Supports matériels et logiciels utilisés.
  - Durée du traitement.
  - Catégories de destinataires et de personnes concernées.

Exemples :

- Une plateforme de vidéo à la demande réutilisant une base de données clients pour entraîner un système de recommandations est responsable de ce nouveau traitement.
- Un fournisseur d'agent conversationnel utilisant des données publiquement accessibles pour entraîner un modèle de langage est responsable de traitement.

2. Réutilisation de données collectées par un autre organisme :

- Diffuseur des données : Met à disposition des données personnelles pour réutilisation.
- Réutilisateur des données : Exploite les données pour son propre compte.
- Chaque acteur est responsable de son propre traitement.

Exemple : Une administration rend publiques des données immobilières ; une entreprise les réutilise pour développer un système d'IA de prédiction immobilière. Les deux sont responsables de traitements distincts.

3. Responsables conjoints du traitement :

Plusieurs responsables de traitement déterminant conjointement les finalités et les moyens du traitement.

Exemples :

- Des centres hospitaliers universitaires utilisant un protocole d'apprentissage fédéré pour analyser des données d'imagerie médicale.
- Un consortium (commune, sociétés de logiciels et de dispositifs vidéo) collaborant pour expérimenter un système de caméras augmentées.

Obligations :

- Définir clairement les rôles et responsabilités dans un accord.
- Chaque responsable peut être contacté par les personnes concernées pour exercer leurs droits.

4. Sous-traitant :

Traite des données pour le compte du responsable de traitement, selon les instructions documentées.

Exemples :

- Un prestataire chargé de constituer une base de données d'apprentissage pour un fournisseur de système d'IA agit en tant que sous-traitant.
- Un prestataire qui crée une base de données à son initiative et développe des systèmes d'IA pour divers clients est responsable de traitement pour cette base de données.

Obligation : Le fournisseur de système d'IA doit s'assurer que le sous-traitant respecte le RGPD et limite le traitement aux instructions données.

IV. Conclusion

La détermination de la qualification juridique des acteurs dans le cadre du RGPD nécessite une analyse détaillée des objectifs et moyens du traitement de données. Les responsabilités peuvent varier entre responsable de traitement, responsable conjoint et sous-traitant, chaque rôle impliquant des obligations spécifiques en matière de protection des données personnelles.

4

LES FICHES PRATIQUES IA

ASSURER QUE LE TRAITEMENT EST LICITE

4.1 DÉFINIR UNE BASE LÉGALE

I. Introduction

L'organisme souhaitant constituer une base de données d'apprentissage contenant des données personnelles pour entraîner un algorithme doit s'assurer que le traitement est autorisé par la loi. La CNIL fournit des directives pour définir les bases légales appropriées selon les responsabilités et les modalités de collecte ou de réutilisation des données.

II. Bases légales du traitement

La constitution et l'utilisation d'une base de données pour l'entraînement de modèles d'IA contenant des données personnelles doivent reposer sur l'une des bases légales prévues par le RGPD.

1. Consentement

Critères de validité : libre, spécifique, éclairé et univoque.

Exemples :

- Filmer ou photographier des volontaires pour entraîner un système à détecter des gestes spécifiques, avec leur consentement.
- Le consentement initial pour une finalité (comme la communication) ne couvre pas la réutilisation pour une autre (comme l'apprentissage d'IA), nécessitant deux consentements distincts.

2. Intérêt Légitime

Conditions :

- Légitimité de l’intérêt poursuivi.
- Nécessité du traitement de données.
- Absence d’atteinte disproportionnée aux droits et intérêts des personnes concernées.

Exemples :

- Collecte de commentaires en ligne pour développer un système d’IA évaluant des œuvres d’art. L'intérêt à développer et commercialiser ce système peut être considéré comme légitime.
- Ne peut pas s'appliquer si l’intérêt commercial porte atteinte aux droits des personnes, comme prédire le profil psychologique à partir de données en ligne.

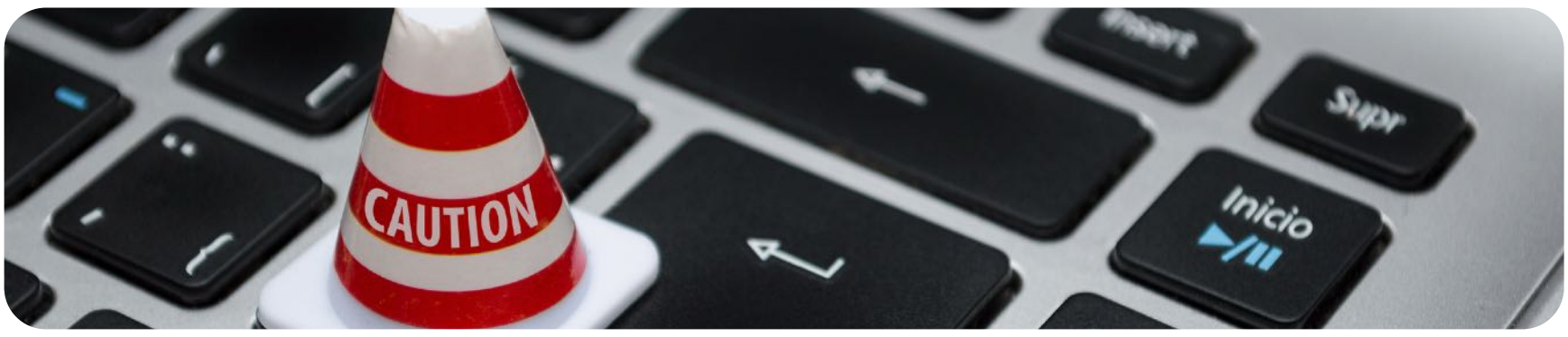
3. Mission d’intérêt public

Conditions :

- La mission doit être prévue par un texte normatif applicable au responsable du traitement.
- l’utilisation des données doit permettre d’exercer cette mission de manière appropriée.

Exemples :

- Analyse de l’évolution de l’utilisation de la langue en ligne par un laboratoire de recherche public.
- Utilisation des données par le PEReN pour réguler les opérateurs de plateformes en ligne.



4. Contrat

Conditions :

- Un contrat valide entre le responsable et la personne concernée.
- Le traitement doit être objectivement nécessaire à l'exécution du contrat.

Exemples :

- Génération automatisée et personnalisée de courriers par un logiciel, nécessitant les données des utilisateurs.
- Ne peut pas être utilisé pour des traitements non indispensables à l'exécution du service, comme l'amélioration de nouveaux produits par un réseau social.

5. Obligation légale

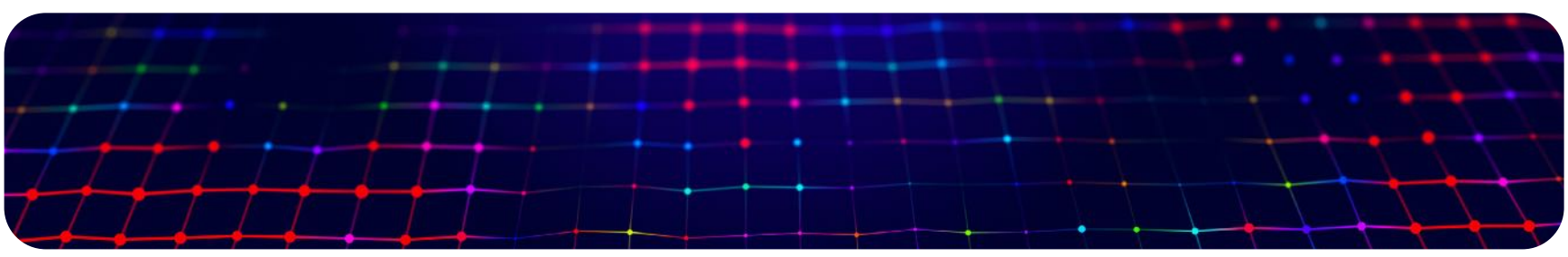
Conditions :

- Le traitement doit être nécessaire pour répondre à une obligation légale déterminée.
- La finalité du traitement doit être définie par un texte normatif.

Exemple : Développement de systèmes d'IA pour les études actuarielles dans le secteur de l'assurance, où l'intérêt légitime pourrait être plus approprié si l'obligation légale n'est pas assez précise.

6. Sauvegarde des intérêts vitaux

Le traitement est nécessaire à la sauvegarde des intérêts vitaux de la personne concernée ou d'une autre personne physique (article 6.1.d) du RGPD).



Exemple : Une personne inconsciente est admise à l'hôpital et ne peut pas donner son consentement pour le traitement de ses données en vue de son traitement.

III. Conclusion

La détermination de la base légale pour constituer une base de données d'apprentissage est essentielle pour assurer la conformité au RGPD. Chaque base légale implique des obligations spécifiques pour le responsable de traitement et doit être choisie en fonction des caractéristiques précises du traitement envisagé. Les exemples pratiques fournis aident à comprendre comment appliquer ces bases légales dans divers contextes de développement de systèmes d'IA.

4

LES FICHES PRATIQUES IA

ASSURER QUE LE TRAITEMENT EST LICITE

4.2 EN CAS DE RÉUTILISATION DES DONNÉES, EFFECTUER LES TESTS ET VÉRIFICATIONS NÉCESSAIRES

I. Introduction

Pour constituer une base de données d'apprentissage contenant des données personnelles, il est essentiel de s'assurer que le traitement de ces données est conforme à la loi. La CNIL fournit des directives pour déterminer les obligations légales en fonction des modalités de collecte et des sources des données.

II. Principes généraux

Le responsable du traitement doit vérifier que le traitement des données est autorisé par la loi, particulièrement en cas de réutilisation de données collectées à l'origine pour une autre finalité. Cela inclut la réalisation de tests de compatibilité pour s'assurer que les nouvelles finalités sont compatibles avec les finalités initiales.

III. Types de réutilisation des données et vérifications

1. Réutilisation des données collectées initialement pour une autre finalité

Obligation de test de compatibilité :

- Nécessaire pour les traitements ultérieurs qui ne s'appuient pas sur le consentement ou sur une obligation légale.

- Doit prendre en compte :
  - Le lien entre la finalité initiale et la finalité ultérieure.
  - Le contexte de collecte des données.
  - La nature et la sensibilité des données.
  - Les conséquences du traitement ultérieur pour les personnes concernées.
  - L'existence de garanties appropriées (pseudonymisation, chiffrement).

Exemples :

- Réutilisation des corrections manuelles apportées par les utilisateurs à des textes générés automatiquement pour personnaliser un service.
- Utilisation de l'historique et des listes de lecture d'une plateforme de streaming pour améliorer les recommandations.

2. Réutilisation des données publiquement accessibles

Conditions :

- Le réutilisateur doit vérifier que la base de données respecte le RGPD et d'autres règles pertinentes.
- Doit s'assurer que la constitution ou la diffusion de la base ne résulte pas d'un crime ou d'un délit.
- La base de données ne doit pas contenir de données sensibles ou d'infractions, sauf vérifications supplémentaires.

Exemples :

- Une entreprise souhaitant utiliser des données de géolocalisation sans source précise ne peut pas ignorer les doutes sur la licéité de telles données.
- Réutilisation de données pseudonymisées initialement rendues publiques par les personnes concernées est acceptable.

3. Réutilisation des données acquises auprès de tiers (courtiers en données)

Conditions :

- Le tiers doit s'assurer de la licéité de la transmission des données.
- Si les données ont été collectées pour une autre finalité, un test de compatibilité doit être effectué.

Exemple : Contrat entre un fournisseur de système d'IA générative d'image et un courtier en données pour garantir la licéité des données partagées.

IV. Obligations des réutilisateurs

Vérifications à effectuer :

- S'assurer que la base de données est licite et conforme au RGPD.
- Informer les personnes concernées et leur permettre d'exercer leurs droits.
- Respecter les exigences de transparence et de documentation de la conformité.

Vérifications spécifiques en cas de données sensibles ou de transmission à des tiers :

- S'assurer de l'obtention du consentement explicite pour les données sensibles.
- Vérifier la licéité des partages de données et documenter les garanties fournies par les tiers.

V. Conclusion

Le respect des obligations légales lors de la réutilisation des données est crucial pour assurer la conformité avec le RGPD. Les responsables de traitement doivent effectuer des vérifications rigoureuses et documenter chaque étape pour garantir la légitimité de leurs actions. La CNIL propose des modèles et des guides pour faciliter ces démarches.

5

LES FICHES PRATIQUES IA

RÉALISER UNE ANALYSE D'IMPACT

I. Introduction

La constitution d'une base de données pour l'apprentissage d'un système d'intelligence artificielle (IA) peut entraîner des risques élevés pour les droits et libertés des personnes concernées. Dans ces cas, une **analyse d'impact sur la protection des données (AIPD)** est obligatoire. Ce document explique comment et dans quels cas réaliser cette analyse.

II. Principes de l'AIPD

L'AIPD permet de :

- Cartographier et évaluer les risques d'un traitement sur la protection des données personnelles.
- Établir un plan d'action pour réduire ces risques à un niveau acceptable.
- Assurer un suivi dans le temps des risques identifiés.

III. Quand une AIPD est-elle nécessaire ?

L'AIPD est obligatoire si le traitement envisagé est susceptible d'engendrer un risque élevé pour les droits et libertés des personnes physiques, conformément à l'article 35 du RGPD.

LES FICHES PRATIQUES IA

RÉALISER UNE ANALYSE D'IMPACT

Le Comité européen de la protection des données (CEPD) a identifié neuf critères pour déterminer si une AIPD est requise. Tout traitement remplissant au moins deux de ces critères nécessitera une AIPD. Ces critères incluent :

- la collecte de données sensibles ou hautement personnelles
- la collecte de données à grande échelle
- la collecte de données de personnes vulnérables (comme les mineurs)
- le croisement ou la combinaison d'ensembles de données
- l'utilisation innovante ou de nouvelles solutions technologiques

IV. Évaluation de l'AIPD pour le développement des systèmes d'IA

1. Usage innovant :
   - Systèmes utilisant des techniques d'IA nouvelles, comme l'apprentissage profond (Deep Learning), nécessitent souvent une AIPD.
   - Systèmes utilisant des techniques éprouvées, comme les forêts aléatoires, peuvent ne pas nécessiter une AIPD.

2. Traitement à grande échelle :
   - Déterminer si le développement concerne un très grand nombre de personnes.
   - Par exemple, une base de données de millions d'images comportant des individus reconnaissables nécessitera une AIPD.

V. Critères de risque et périmètre de l'AIPD

1. Réglementation européenne sur l'IA : Les systèmes d'IA à haut risque, comme définis par le projet de règlement européen sur l'IA, nécessitent une AIPD.

2. Modèles de fondation et systèmes à usage général : nécessitent souvent une AIPD en raison de la diversité des usages possibles et des risques associés.

3. Définir le Périmètre de l'AIPD :
Si l'usage opérationnel du système est identifié dès la phase de développement, une AIPD générale peut être réalisée.
Si l'usage opérationnel n'est pas identifié, l'AIPD doit être modifiée de manière itérative au fur et à mesure de l'évolution du traitement.

VI. Risques spécifiques à l'IA à prendre en compte dans l'AIPD

- Risques de mésusage des données, de violation de données et de discrimination automatisée.
- Risques de contenu fictif erroné, de prise de décision automatisée biaisée.
- Risques d'attaques spécifiques aux systèmes d'IA, comme les attaques par empoisonnement des données.
- Risques éthiques et de confidentialité des données.

VII. Mesures à prendre en fonction des résultats de l'AIPD

1. Mesures techniques :
- Sécurité : chiffrement homomorphe, environnement d'exécution sécurisé.
- Minimisation : recours à des données synthétiques.
- Anonymisation/Pseudonymisation : confidentialité différentielle.
- Protection des données dès le développement : apprentissage fédéré.
- Facilitation des droits : techniques de désapprentissage machine, explicabilité et traçabilité.

2. Mesures organisationnelles et de gouvernance :
- Limitation de l'accès aux bases de données.
- Mise en place d'un comité éthique.
- Documentation interne et traçabilité des actions.

LES FICHES PRATIQUES IA

RÉALISER UNE ANALYSE D'IMPACT

VIII. Conclusion

La réalisation d'une AIPD est cruciale pour assurer la conformité des systèmes d'IA avec le RGPD. Elle permet de minimiser les risques pour les droits et libertés des personnes concernées en intégrant des mesures techniques et organisationnelles appropriées. La publication partielle de l'AIPD est recommandée pour garantir la transparence et la confiance du public.

6

LES FICHES PRATIQUES IA

TENIR COMPTE DE LA PROTECTION DES DONNÉES DANS LA CONCEPTION DU SYSTÈME

I. Introduction

Pour assurer le développement d'un système d'IA respectueux de la protection des données, il est nécessaire de mener une réflexion préalable lors de la conception du système.

II. Étapes clés de la conception respectueuse de la protection des données

1. L'objectif du système

Préciser l'emploi du système d'IA : Définir le type de résultats attendus, les indicateurs de performance (quantitatifs comme le F1-score ou qualitatifs provenant de retours humains), et le contexte d'utilisation du système.

Exemple : Pour le comptage des personnes debout dans un tramway, un réseau de neurones détectant la présence de personnes sans analyse de posture pourrait être préféré à un réseau nécessitant plus de données pour analyser la posture.

2. La méthode à employer

Choisir la technique la plus respectueuse : Sélectionner la méthode qui minimise l'utilisation des données tout en atteignant les objectifs. Privilégier les techniques ne nécessitant pas l'apprentissage machine si possible.

Exemples :

- Utiliser un détecteur de mouvement infrarouge plutôt qu'une caméra pour délimiter une zone dangereuse.
- Analyser le contenu textuel par un algorithme moins gourmand en données.

3. Sélectionner les données strictement nécessaires

Appliquer le principe de minimisation : Les données personnelles doivent être adéquates, pertinentes et limitées à ce qui est nécessaire pour les finalités du traitement.

Prendre en compte quatre dimensions :

- Volume : Justifier le nombre de personnes concernées, la profondeur historique, etc.
- Catégories : Examiner et justifier la présence de données sensibles.
- Typologie : Utiliser des données réelles, de synthèse, anonymisées, etc.
- Sources : Recenser les sources de données (collecte initiale ou réutilisation).

4. Valider les choix de conception

Mener une étude pilote : Valider les choix techniques et de données par une expérimentation à petite échelle avec des données fictives ou anonymisées.

Exemples :

- Utiliser des données de réseaux sociaux avec consentement pour étudier le ciblage publicitaire.
- Collecter des données anonymisées pour tester un système de recommandation de films.

5. Interroger un comité éthique

Missions du comité éthique : Formuler des avis, animer des réflexions éthiques, élaborer une doctrine interne, et recommander des principes de comportement.

Constitution du comité :

- Pluridisciplinaire : Diversité des membres en termes de profil, sexe, âge et origines culturelles.
- Indépendant : Les membres ne doivent pas avoir d’intérêt personnel dans les décisions.
- Rôle défini : Procédure claire pour la réunion et l'association du comité, pouvant être consultatif ou contraignant.
- Averti : Se tenir informé des évolutions techniques et des bonnes pratiques.

Questions pour le comité éthique :

- Validité des données utilisées.
- Conséquences des usages opérationnels.
- Risques d’utilisations détournées.
- Maîtrise des choix techniques.
- Suffisance des mesures de transparence.
- Prévention des discriminations.

III. Conclusion

La conception d’un système d’IA respectueux de la protection des données repose sur une série de réflexions et de validations techniques et éthiques. Les étapes clés incluent la définition claire des objectifs, le choix des méthodes et des données adéquates, la validation des choix par des études pilotes, et la consultation de comités éthiques. Ces mesures assurent que les systèmes d’IA développés minimisent les risques pour les droits et libertés des personnes concernées.

7

LES FICHES PRATIQUES IA

TENIR COMPTE DE LA PROTECTION DES DONNÉES DANS LA COLLECTE ET LA GESTION DES DONNÉES

I. Introduction

Le développement d'un système d'intelligence artificielle nécessite une gestion et un suivi rigoureux des données d'apprentissage. La CNIL détaille comment les principes relatifs à la protection des données s'articulent avec la gestion des données d'apprentissage, en intégrant dès leur conception les principes de protection des données personnelles (« privacy by design »).

II. Collecte des données

1. Critères de collecte :

- Limiter la collecte aux données librement accessibles.
- Définir des critères précis de collecte avant la mise en œuvre du traitement.
- Assurer la collecte de données pertinentes uniquement et supprimer immédiatement les données inutiles.

2. Moissonnage de données (« web scraping ») :

- Minimiser la collecte de données.
- Respecter les guides de la CNIL sur l'ouverture, le partage et la réutilisation des données publiquement accessibles.

LES FICHES PRATIQUES IA

TENIR COMPTE DE LA PROTECTION DES DONNÉES DANS LA COLLECTE ET LA GESTION DES DONNÉES

III. Nettoyage, identification et protection de la vie privée dès la conception

1. Nettoyage des données : Corriger les valeurs vides, détecter les valeurs aberrantes, corriger les erreurs, éliminer les doublons, et supprimer les champs inutiles.

2. Identification des données pertinentes : Les données peuvent être brutes (audio, image, texte manuscrit) ou structurées (mesures, observations numériques).

- Les métadonnées fournissent des informations de description, structure et qualité des données.
- Les annotations et caractéristiques extraites de données permettent une sélection des données pertinentes.

3. Techniques de sélection des caractéristiques :

- Utilisation de bibliothèques comme Yellowbrick, LOFO, ou Facets pour la sélection de caractéristiques.
- Approches d'annotations interactives (« active learning ») et techniques d'ablation des données d'entraînement (« data pruning »).

IV. Protection des données dès la conception (« Privacy by Design »)

1. Mesures techniques :

- Généralisation : Modifier l'échelle ou l'ordre de grandeur des attributs.
- Randomisation : Ajouter du bruit aux données pour diminuer la précision et affaiblir le lien avec l'individu.

2. Techniques de cryptographie : Cryptage homomorphe et calcul multipartite sécurisé pour entraîner des modèles d'IA sur des données chiffrées.

3. Mesures spécifiques : Confidentialité différentielle et apprentissage fédéré pour protéger les données durant l'apprentissage du modèle.

V. Suivi et mise à jour

1. Surveillance de la dérive des données :
   - Comparaison régulière des données avec les données sources pour détecter la dérive.
   - Utilisation d'outils comme Evidently et Scipy pour automatiser cette détection.

2. Revue régulière :
   - Analyse périodique par des agents formés aux questions de protection des données ou par un comité d'éthique.
   - Veille scientifique pour identifier de nouvelles techniques de gestion de données.

VI. Conservation des données

1. Durée de conservation :
   - Définir une durée de conservation en fonction de l'objectif ayant conduit à la collecte des données.
   - Planification en amont et suivi de la durée de conservation, incluant la suppression automatique des données après cette durée.

2. Maintenance et amélioration du produit :
   - Conserver uniquement les données nécessaires pour la maintenance ou l'amélioration du système.
   - Utilisation de mesures de sécurité renforcées pour les données conservées.

VII. Sécurité

Mesures de sécurité : Chiffrement des flux et des sauvegardes, vérification de l'intégrité des données, journalisation des opérations, et formation des agents.

7

LES FICHES PRATIQUES IA

TENIR COMPTE DE LA PROTECTION DES DONNÉES DANS LA COLLECTE ET LA GESTION DES DONNÉES

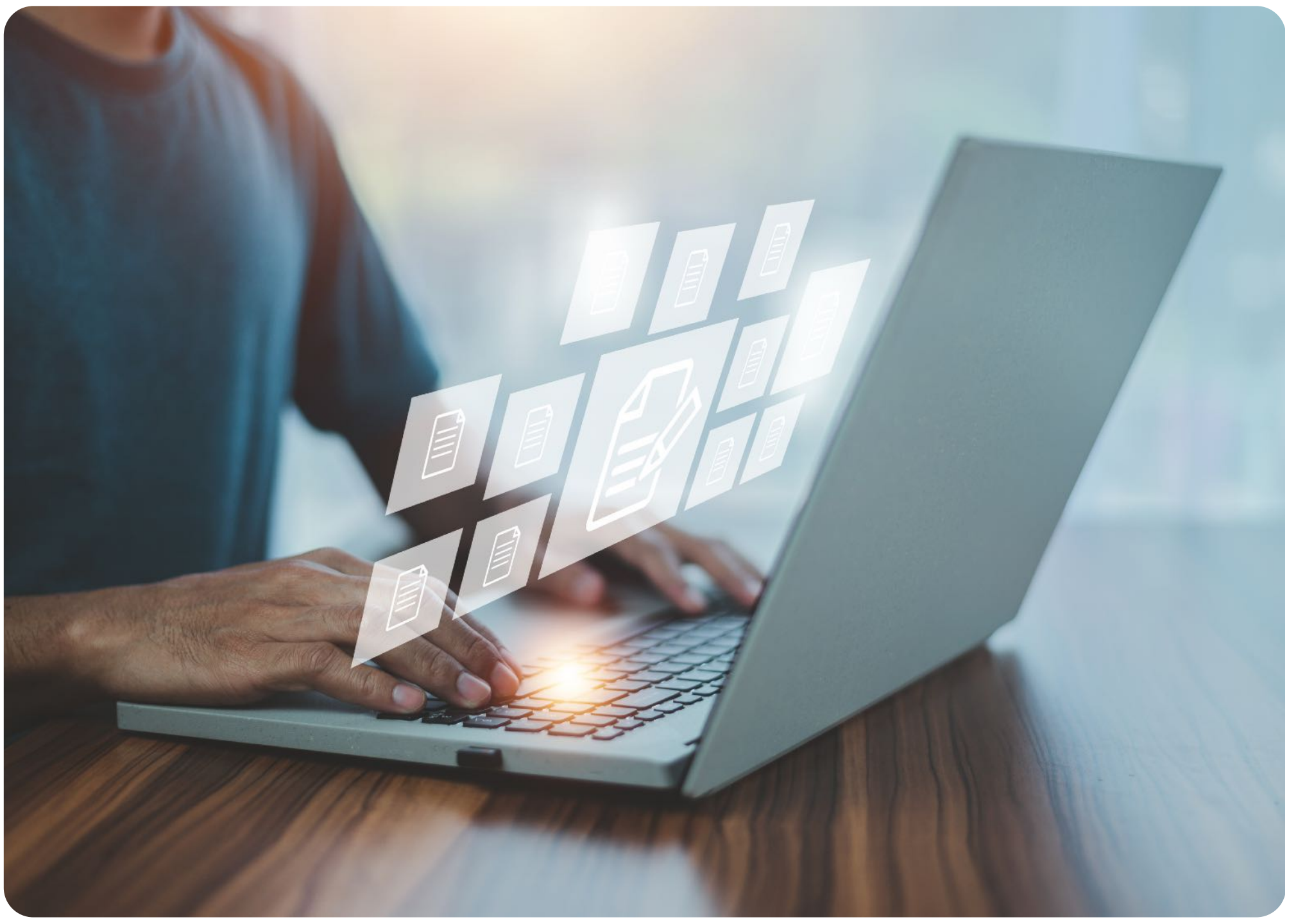
VIII. Documentation

1. Objectifs de la documentation : Assurer la traçabilité des jeux de données, démontrer la licéité de la collecte, faciliter le suivi des données, réduire les risques d'utilisation imprévue, permettre l'exercice des droits des personnes, et identifier les améliorations prévues.

2. Modèles de documentation :

- Utilisation de modèles existants (comme ceux de Gebru et al., 2021) pour structurer la documentation.
- Documentation à fournir aux utilisateurs pour assurer la transparence.

Cette synthèse présente les étapes essentielles pour intégrer la protection des données dès la conception d'un système d'IA, en respectant les principes de minimisation, de sécurité et de documentation rigoureuse.



8

LES FICHES PRATIQUES IA

MOBILISER LA BASE LÉGALE DE L'INTÉRÊT LÉGITIME POUR DÉVELOPPER UN SYSTÈME D'IA

I. Introduction

Le recours à l'intérêt légitime est une base légale courante pour le développement de systèmes d'intelligence artificielle (IA) par des organismes privés. Cette base légale, prévue par l'article 6 du RGPD, est adaptée lorsqu'il est difficile de collecter le consentement des personnes concernées. Les organismes publics peuvent également utiliser cette base légale sous certaines conditions.

II. Conditions de l'intérêt légitime

Pour que l'intérêt légitime soit une base légale valide, trois conditions doivent être respectées :

1. Intérêt légitime : l'intérêt doit être licite, clair et précis, réel et présent pour l'organisme. Exemples d'intérêts légitimes :
   - Travaux de recherche scientifique.
   - Facilitation de l'accès du public à certaines informations.
   - Amélioration de produits ou services.
   - Détection de comportements frauduleux.

2. Nécessité du traitement :
   - Le traitement doit être nécessaire pour atteindre l'objectif poursuivi.
   - L'organisme doit s'assurer qu'il n'existe pas de moyens moins intrusifs pour atteindre cet objectif.

LES FICHES PRATIQUES IA

MOBILISER LA BASE LÉGALE DE L'INTÉRÊT LÉGITIME POUR DÉVELOPPER UN SYSTÈME D'IA

3.Mise en balance des droits et intérêts :

- Le traitement ne doit pas porter une atteinte disproportionnée aux droits et intérêts des personnes concernées.
- L'organisme doit évaluer les avantages du traitement et les impacts sur les individus, et mettre en place des mesures additionnelles pour limiter ces risques.

III. Évaluation des impacts et des bénéfices

1. Bénéfices pour l'organisme et les tiers : Amélioration des soins de santé, meilleure accessibilité des services, facilitation de l'exercice des droits fondamentaux.

Exemple : Un système de reconnaissance vocale pour aider les personnes en situation de handicap.

2. Utilité pour la conformité réglementaire : Par exemple, un système d'IA pour la modération de contenu en ligne conformément aux dispositions de l'article 35.1 du DSA.

3. Transparence et partage des bénéfices : Développement en source ouverte pour bénéficier à la communauté scientifique.

IV. Identification des risques pour les personnes concernée

1. Collecte des données :

- Risques d'atteinte à la vie privée, collecte illégale, atteinte à la liberté d'expression.
- Attention particulière aux données sensibles et aux personnes vulnérables.

2. Entraînement du modèle et conservation des données : Risques de perte de confidentialité, difficulté à garantir l'exercice des droits, manque de transparence.

LES FICHES PRATIQUES IA

MOBILISER LA BASE LÉGALE DE L'INTÉRÊT LÉGITIME POUR DÉVELOPPER UN SYSTÈME D'IA

3.Utilisation du Système d'IA : Risques de mémorisation et régurgitation de données personnelles, atteinte à la réputation, risques éthiques graves.

V. Mesures de protection additionnelles

1. Collecte et constitution de la base de données :
- Anonymisation ou pseudonymisation des données.
- Utilisation de données synthétiques pour minimiser les risques.

2. Entraînement du modèle et conservation des données :
- Facilitation de l'exercice des droits (droit d'opposition, transparence accrue).
- Développement transparent du système (documentation, journalisation des activités).

3. Utilisation du système d'IA
- Limitation des risques de régurgitation et génération de données personnelles.
- Prévention des biais discriminatoires et réutilisation illicite.

VI. Conclusion

Mobiliser la base légale de l'intérêt légitime pour le développement de systèmes d'IA nécessite une évaluation rigoureuse des intérêts et des risques. Les organismes doivent mettre en place des mesures techniques, organisationnelles et juridiques pour assurer la protection des droits et libertés des personnes concernées tout en maximisant les bénéfices de l'utilisation des systèmes d'IA.

8

LES FICHES PRATIQUES IA

MOBILISER LA BASE LÉGALE DE L'INTÉRÊT LÉGITIME POUR DÉVELOPPER UN SYSTÈME D'IA

8.1 FOCUS SUR LA DIFFUSION DES MODÈLES EN SOURCE OUVERTE (OPEN SOURCE)

I. Introduction

La diffusion des modèles d'intelligence artificielle (IA) en source ouverte présente des bénéfices significatifs pour la transparence, le contrôle par les pairs, et l'innovation scientifique. Toutefois, elle doit être encadrée pour limiter les atteintes aux droits des personnes concernées. Cette fiche détaille les pratiques, les bénéfices et les risques liés à l'ouverture des modèles en source ouverte, ainsi que les garanties nécessaires à mettre en place.

II. Pratiques d'ouverture en IA

1. Transparence sur le développement du modèle :

- Publication de la documentation de la procédure suivie (y compris la collecte des données d'apprentissage), éventuellement sous forme de publication scientifique.
- Publication du code utilisé pour entraîner le modèle.
- Publication des données d'entraînement (sous conditions).

2. Transparence sur le modèle obtenu :

- Documentation détaillée du modèle (architecture, performances, limitations) sous forme de fiche descriptive (model card).
- Publication des poids du modèle.

3. Accès au modèle :

- Bibliothèque, API, ou code permettant l'utilisation du modèle.
- Modèle diffusé sous une licence permettant son utilisation, modification, ou redistribution.

III. Bénéfices de l'ouverture en IA

1. Transparence et exercice des droits :
   - Augmentation de la transparence du modèle et de son fonctionnement.
   - Vérification des capacités, limitations, et présence de biais dans le modèle.

2. Sécurité et qualité du modèle :
   - Détection et correction des vulnérabilités du modèle.
   - Amélioration de la sécurité et réduction des biais.

3. Innovation et recherche scientifique :
   - Partage de connaissances entre développeurs.
   - Accessibilité pour les étudiants et chercheurs.
   - Harmonisation des pratiques et interopérabilité des modèles.

4. Bénéfices pour les utilisateurs :
   - Transparence accrue pour les utilisateurs du modèle.
   - Responsabilisation des acteurs grâce au contrôle par les pairs.

IV. Risques de l'ouverture en IA

1. Réutilisation illicite ou malveillante :
   - Multiplicité des réutilisations non maîtrisées pouvant conduire à des utilisations illicites.
   - Risques de contournement des filtres et sécurités ajoutés aux systèmes.

2. Sécurité du modèle :
   - Exploitation des failles du modèle par des attaquants.
   - Risques liés à la mémorisation et régurgitation des données d'entraînement.

3. Confidentialité des données :
   - Publication des données d'entraînement peut porter atteinte à la confidentialité.
   - Difficulté à assurer l'exercice des droits tout au long de la chaîne d'utilisation et de redistribution du modèle.

V. Garanties additionnelles à mettre en place

1. Publication d'éléments clés :
   - Paramètres du modèle.
   - Code nécessaire pour utiliser le modèle.
   - Fiche descriptive du modèle et des données utilisées pour l'apprentissage.

2. Mesures juridiques et techniques :
   - Licences restrictives pour limiter la réutilisation des modèles.
   - Tatouage numérique pour tracer et contrôler les réutilisations.
   - Anonymisation ou pseudonymisation des données pour sécuriser les données d'entraînement.

3. Information et exercice des droits :
   - Information accrue des personnes concernées.
   - Transmission de l'exercice des droits le long de la chaîne des acteurs.
   - Traçabilité des téléchargements pour les cas à risque.

4. Transparence et auditabilité :
   - Documentation et journalisation des activités de développement.
   - Contrôle effectif du développement par les pairs.
   - Mise en place de comités éthiques ou de référents éthiques.

VI. Conclusion

La diffusion en source ouverte de modèles d'IA peut renforcer l'intérêt légitime des responsables de traitement en offrant des bénéfices importants pour la recherche scientifique et la transparence. Cependant, elle doit être accompagnée de garanties appropriées pour limiter les risques potentiels pour les droits et libertés des personnes concernées. Les mesures recommandées par la CNIL permettent de concilier ouverture et protection des données personnelles.

8

LES FICHES PRATIQUES IA

MOBILISER LA BASE LÉGALE DE L'INTÉRÊT LÉGITIME POUR DÉVELOPPER UN SYSTÈME D'IA

8.2 FOCUS SUR LES MESURES À PRENDRE EN CAS DE COLLECTE DES DONNÉES PAR MOISSONNAGE (WEB SCRAPING)

I. Introduction

La collecte de données accessibles en ligne par moissonnage (web scraping) est de plus en plus courante, surtout avec l'essor des systèmes d'IA générative. Cependant, ces pratiques peuvent poser des risques importants pour les droits et libertés des personnes concernées. Cette fiche de la CNIL rappelle les obligations des responsables de traitement et précise les conditions de mise en œuvre de telles pratiques pour le développement d'un système d'IA.

II. Doctrine de la CNIL

- La collecte de données par moissonnage doit être accompagnée de mesures garantissant les droits des personnes concernées.
- Les pratiques de moissonnage reposent souvent sur la base légale de l'intérêt légitime, nécessitant des mesures additionnelles pour limiter les atteintes aux droits des personnes.

III. Mesures obligatoires

1. Critères de collecte précis :
   - Définir des critères précis pour exclure la collecte de données non nécessaires (données bancaires, géolocalisation, mineurs, données sensibles).
   - Supprimer immédiatement les données non pertinentes collectées par erreur.

2. Attention aux données sensibles :

- Exclure automatiquement les données sensibles par des filtres.
- Supprimer les données sensibles collectées de manière incidente dès leur identification.

IV. Garanties supplémentaires

1. Exclusion de sites spécifiques :

- Exclure par défaut la collecte de données provenant de sites contenant des données intrusives ou sensibles (sites pornographiques, forums de santé, réseaux sociaux de mineurs).
- Exclure les sites qui s'opposent au moissonnage via des fichiers robot.txt ou ai.txt.

2. Liste repoussoir et opposition au traitement :

- Mettre en place une liste repoussoir permettant aux personnes de s'opposer à la collecte de leurs données.
- Permettre aux personnes de s'opposer au traitement de leurs données de manière discrétionnaire.

3. Limitation aux données libre d'accès : Limiter la collecte aux données librement accessibles et manifestement rendues publiques par les utilisateurs.

4. Anonymisation et pseudonymisation : Anonymiser ou pseudonymiser les données juste après leur collecte pour minimiser les risques.

5. Information et transparence : Diffuser largement les informations relatives à la collecte et aux droits des personnes, et publier une liste des sites concernés par le moissonnage.

V. Registre des organisations utilisant le moissonnage

1. Projet de registre :
   - La CNIL envisage de mettre en place un registre des organisations traitant des données collectées par moissonnage pour des fins de développement de systèmes d'IA.
   - Ce registre concernerait les organisations publiques et privées utilisant des outils de moissonnage.

2. Informations sur le registre :
   - Identité du responsable du traitement.
   - Lien vers la politique de confidentialité.
   - Liste des catégories de sites moissonnés.
   - Modalités d'exercice des droits.

3. Inscription volontaire :
   - L'inscription sur le registre ne serait pas obligatoire mais favoriserait la transparence et faciliterait l'exercice des droits des personnes.
   - La présence sur le registre ne garantirait pas la conformité des traitements de données mais indiquerait un engagement envers la transparence.

VI. Conclusion

La collecte de données par moissonnage pour le développement de systèmes d'IA doit être encadrée par des mesures strictes pour protéger les droits et libertés des personnes concernées. La CNIL recommande des pratiques spécifiques pour garantir la minimisation des données, la transparence, et l'exercice des droits. L'inscription sur un registre de la CNIL pourrait constituer une mesure supplémentaire de transparence et de protection.

9

LES FICHES PRATIQUES IA

INFORMER LES PERSONNES CONCERNÉES

I. Introduction

Les organismes qui traitent des données personnelles pour développer des modèles ou des systèmes d’intelligence artificielle (IA) doivent informer les personnes concernées. Cette fiche, soumise à consultation publique, précise les obligations en matière de transparence et d'information des personnes.

II. Assurer la transparence des traitements

1. Principe de transparence :
   - Informer les personnes concernées sur les usages de leurs données.
   - Permettre aux personnes d’exercer leurs droits (opposition, accès, rectification, etc.).

2. Quand informer :
   - Lors de la collecte directe : au moment de la collecte.
   - Lors de la collecte indirecte : dès que possible, au plus tard lors de la première communication ou dans un délai d'un mois.

3. Exceptions : Dispense d’information individuelle si elle est impossible ou nécessite des efforts disproportionnés.

LES FICHES PRATIQUES IA

INFORMER LES PERSONNES CONCERNÉES

III. Quelles informations fournir ?

1. Informations générales :
- Identité et coordonnées du responsable du traitement.
- Finalité et base légale du traitement.
- Destinataires ou catégories de destinataires des données.
- Durée de conservation des données.
- Droits des personnes concernées (accès, rectification, effacement, etc.).
- Droit de réclamation auprès de la CNIL.

2. Informations supplémentaires en cas de collecte indirecte :
- Catégories de données personnelles collectées.
- Sources des données.

3. Cas particuliers :
- Moissonnage (web scraping) : détails sur les plateformes et les types de données collectées.
- Réutilisation de données ou modèles : informations sur les sources et modalités de collecte.

IV. Exemples de mentions informatives

1.Données d'entraînement :
- Provenance des données (par ex., jeux de données publiquement accessibles ou fournis par des courtiers de données).
- Informations détaillées sur la collecte et l'annotation des données.

2. Moissonnage : Mention des sites moissonnés et des types de données collectées.

3. Modèles d'IA soumis au RGPD
- Destinataires du modèle.
- Durée de conservation et droits des personnes concernées.
- Nature des risques liés à la régurgitation des données.

LES FICHES PRATIQUES IA

INFORMER LES PERSONNES CONCERNÉES

V. Quand et comment fournir l'information ?

1. Moment de l'Information :
   - Collecte directe : informer au moment de la collecte.
   - Collecte indirecte : informer dès que possible, au plus tard dans un délai d'un mois.

2. Moyens de Fournir l'Information :
   - Information individuelle : via formulaires en ligne, emails, courriers, etc.
   - Information générale : via sites web, panneaux d'affichage, etc.

VI. Garantir l'accessibilité et l'intelligibilité de l'information

1. Accessibilité : Mentions d'information doivent être facilement accessibles et distinguées des autres informations légales.

2. Intelligibilité :
   - Informations succinctes et claires, adaptées aux conditions d'interaction avec les personnes.
   - Utilisation de schémas pour expliquer le traitement des données et le fonctionnement du système d'IA.
   - Information en plusieurs niveaux pour prioriser les informations essentielles.

VII. Cas d'exceptions à l'obligation d'informer individuellement

1. Dérogations prévue par le RGPD :
   - Personne concernée déjà informée.
   - Information individuelle disproportionnée (ancienneté des données, absence de coordonnées, etc.).

2. Bonnes pratiques en cas de collecte indirecte :
   - S'appuyer sur le diffuseur de données pour informer les personnes.
   - Publier des informations sur le site internet de l'organisme.

LES FICHES PRATIQUES IA

INFORMER LES PERSONNES CONCERNÉES

VIII. Bonnes pratiques pour la transparence

1.Publication de l'AIPD : Partielle ou complète, selon les sections soumises à des secrets protégés.

2. Documentation du développement : Documentation sur la base de données, le processus de développement et le système d'IA.

3. Pratiques de transparence :

- Publication des poids du modèle, du code source.
- Transparence sur les concepts clés de l'IA et les mesures de sécurité.

IX. Conclusion

Informer les personnes concernées est essentiel pour garantir la transparence des traitements de données dans le développement des systèmes d'IA. Les organismes doivent fournir des informations complètes, accessibles et intelligibles, en respectant les obligations du RGPD et en adoptant des bonnes pratiques pour renforcer la confiance des personnes dans l'utilisation de leurs données.

10

LES FICHES PRATIQUES IA

RESPECTER ET FACILITER L'EXERCICE DES DROITS DES PERSONNES CONCERNÉES

I. Introduction

Les personnes dont les données sont collectées, utilisées ou réutilisées pour développer un système d'IA disposent de droits sur leurs données. Il appartient aux responsables de traitement de respecter ces droits et de faciliter leur exercice. Cette fiche précise les droits applicables et les mesures à mettre en œuvre pour leur respect.

II. Droits applicables

Les droits des personnes concernées incluent :

1. Droit d'accès (article 15 du RGPD) : Connaître les données qu'un organisme détient sur elles.

2. Droit de rectification (article 16 du RGPD) : Corriger les informations incorrectes.

3. Droit à l'effacement ou droit à l'oubli (article 17 du RGPD) : Supprimer les données personnelles sous certaines conditions.

4. Droit à la limitation du traitement (article 18 du RGPD) : Restreindre le traitement des données.

5. Droit à la portabilité des données (article 20 du RGPD) : Obtenir et réutiliser une copie des données.

6. Droit d'opposition (article 21 du RGPD) : Refuser l'utilisation des données pour des raisons tenant à leur situation particulière.

7. Droit de retirer son consentement (article 7.3 du RGPD) : Retirer son consentement à tout moment.

III. Exercice des droits sur les bases de données d'apprentissage

1. Identification des personnes concernées :
   - Le responsable du traitement ne doit pas conserver des informations supplémentaires uniquement pour permettre l'exercice des droits si elles ne sont plus nécessaires.
   - Les personnes peuvent fournir des informations complémentaires pour leur identification (ex. : image, pseudonyme).

2. Informations sur le traitement de données :
   - Fournir des informations sur les destinataires des données et leur source.
   - Conserver les noms de domaine et les URL des pages web moissonnées.

3. Copie des données d'apprentissage :
   - Droit de recevoir une copie des données et des annotations dans un format compréhensible.
   - Ne pas porter atteinte aux droits et libertés d'autrui.

4. Modification et notification :
   - Droit de compléter ou rectifier les données.
   - Droit à l'effacement dans certains cas (données sensibles, par exemple).
   - Notification des rectifications, limitations ou effacements aux destinataires des données.

LES FICHES PRATIQUES IA

RESPECTER ET FACILITER L'EXERCICE DES DROITS DES PERSONNES CONCERNÉES

IV. Exercice des droits sur les modèles d'IA

1. Identification dans le modèle :
- Le responsable peut démontrer qu'il n'est pas en mesure d'identifier les personnes dans le modèle.
- Utilisation de techniques comme les fonctions d'influence pour identifier la contribution des données.

2. Informations sur le modèle :
- Fournir des informations spécifiques au modèle (destinataires, durée de conservation, droits applicables).
- Réentraînement du modèle pour prendre en compte les demandes d'exercice de droit.

3. Mesures techniques :
- Techniques de désapprentissage machine pour respecter les droits.
- Filtres empêchant la régurgitation de données personnelles.

V. Dérogations à l'exercice des droits

1. Situations de dérogation :
- Demande manifestement infondée ou excessive.
- L'organisme qui reçoit la demande n'est pas le responsable du traitement.
- Droit exclu par le droit français ou européen.
- Risques pour la recherche scientifique ou historique.

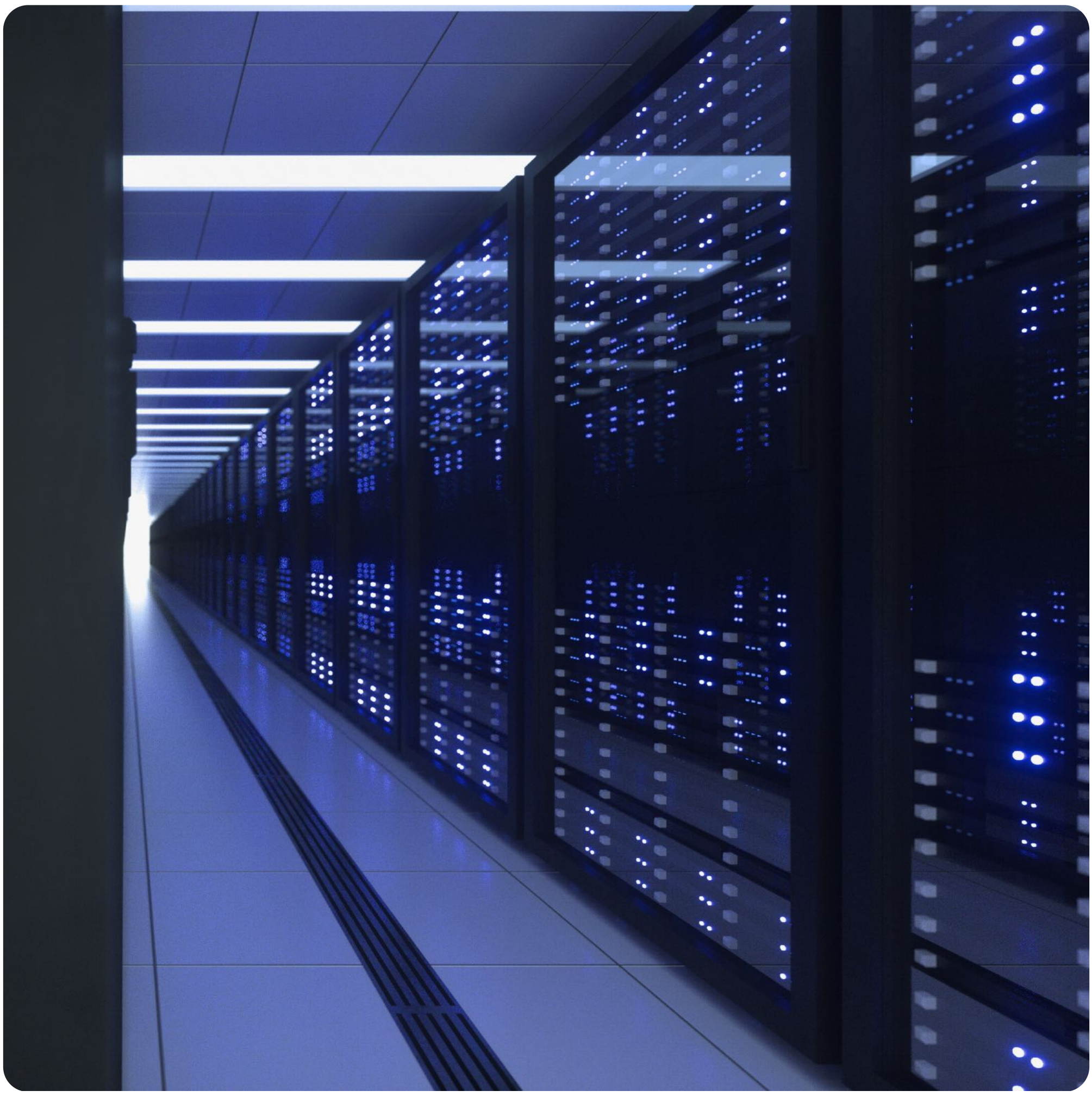
2. Droit d'opposition :
- Mise en balance des motifs légitimes du traitement et des motifs de la personne concernée.
- Appréciation des risques de divulgation des données d'entraînement.

LES FICHES PRATIQUES IA

RESPECTER ET FACILITER L'EXERCICE DES DROITS DES PERSONNES CONCERNÉES

VI. Conclusion

Le respect des droits des personnes dans le cadre du développement des systèmes d'IA est essentiel pour garantir la conformité au RGPD et la protection des données personnelles. Les responsables de traitement doivent mettre en place des mesures techniques et organisationnelles pour faciliter l'exercice de ces droits, tout en tenant compte des défis spécifiques liés aux modèles d'IA.



11

LES FICHES PRATIQUES IA

ANNOTER LES DONNÉES

I. Introduction

L'annotation des données est une étape cruciale dans le développement de modèles d'intelligence artificielle (IA). Cette fiche détaille les bonnes pratiques et les obligations à respecter pour garantir la qualité des annotations tout en protégeant les données personnelles.

II. Importance de l'annotation des données

1. Performance du modèle : Une annotation précise et correcte est essentielle pour obtenir des modèles performants.

2. Respect des droits des personnes : Les annotations doivent respecter les principes de protection des données personnelles, notamment la minimisation, l'exactitude et la loyauté.

III. Exemples d’annotations

- Reconnaissance vocale : enregistrements vocaux annotés avec l’identité du locuteur.
- Détection de chutes : images annotées avec la position des personnes (debout, couché).
- Reconnaissance de plaques minéralogiques : images annotées avec la position des pixels contenant une plaque.
- Prédiction de pathologies : résultats sanguins annotés avec le diagnostic médical.

IV. Principes à respecter

1. Minimisation
- Ne traiter que les données nécessaires.
- Limiter les annotations à ce qui est pertinent pour la finalité du traitement.

2. Exactitude :
- Les annotations doivent être exactes et, si nécessaire, mises à jour.
- Utiliser des critères d’annotation objectifs et éviter les informations ambiguës ou inexactes.

3. Loyauté et transparence : Les personnes doivent être informées des annotations réalisées sur leurs données.

V. Protocole d'annotation

1.Choix des labels :
- Les labels doivent être adaptés à la finalité visée et permettre une annotation objective.
- Exclure tout terme dégradant ou discriminant.

2. Procédure d'annotation :
- Documenter la procédure.
- Limiter l’accès aux données aux personnes habilitées.
- Prévoir une phase de validation pour confirmer le choix des labels et le fonctionnement de la procédure.

3. Vérification continue :
- Mettre en place une procédure de contrôle qualité dès le début de l’annotation.
- Réaliser des contrôles réguliers pour s’assurer de la pertinence et de l’exactitude des annotations.

VI. Information et exercice des droits

1. Informer les personnes concernées :
   - Expliquer l'objectif de l'annotation, l'organisme responsable et les mesures de sécurité mises en place.
   - Informer des résultats de l'annotation lorsque cela est possible.

2. Exercice des droits :
   - Les annotations sont soumises aux droits d'accès, de rectification, d'effacement, d'opposition et de limitation.
   - Les annotations doivent être corrigées ou supprimées en fonction des demandes des personnes concernées.

VII. Annotation des données sensibles

1. Données sensibles :
   - Les annotations révélant des données sensibles doivent respecter les exceptions prévues par le RGPD.
   - Prendre des mesures particulières pour limiter les risques, telles que l'utilisation de données de synthèse.

2. Mesures de sécurité :
   - Réaliser l'annotation en interne ou sécuriser les données si traitées en externe.
   - Utiliser des techniques d'annotation objectives pour éviter les biais et les discriminations.

VIII. Conclusion

L'annotation des données doit être réalisée avec rigueur et en respectant les principes de protection des données personnelles. Les responsables de traitement doivent mettre en place des procédures robustes et des mesures de sécurité pour garantir la qualité des annotations et protéger les droits des personnes concernées.

12

LES FICHES PRATIQUES IA

GARANTIR LA SÉCURITÉ DU DÉVELOPPEMENT D'UN SYSTÈME D'IA

I. Introduction

La sécurité des systèmes d'intelligence artificielle (IA) est cruciale pour garantir la protection des données personnelles et des droits des personnes concernées. Cette fiche détaille les objectifs, les facteurs de risques, et les mesures recommandées pour sécuriser le développement d'un système d'IA, conformément à l'article 32 du RGPD.

II. Approche méthodologique à adopter

1. Analyse des risques :
   - Réaliser une analyse des risques spécifique aux systèmes d'IA et aux bases de données d'entraînement de grande taille.
   - Prendre en compte les sources de risques, les supports visés, et les impacts sur les personnes.

2. Vigilance sur les composants utilisés : Évaluer la sécurité des bases de données, bibliothèques, et infrastructures utilisées, surtout pour des usages présentant des risques élevés.

3. Analyse d'impact sur la protection des données (AIPD) : Conduire une AIPD pour répertorier les risques et les mesures adaptées, obligatoire dans certains cas.

LES FICHES PRATIQUES IA

GARANTIR LA SÉCURITÉ DU DÉVELOPPEMENT D'UN SYSTÈME D'IA

III. Les objectifs de sécurité liés au développement en IA

1. Confidentialité des données :
   - Prévenir les pertes de confidentialité des données d'entraînement.
   - Prendre en compte les risques de mémorisation de données, de reconstruction, ou d'inférence d'appartenance.

2. Performance et intégrité du système :
   - Assurer la fiabilité des données, outils, et protocoles utilisés.
   - Maintenir les performances du système dans le temps, notamment pour les systèmes d'apprentissage continu.

3. Sécurité générale du système d'information : Protéger les autres composants du système (sauvegardes, interfaces, communications) pour prévenir les intrusions.

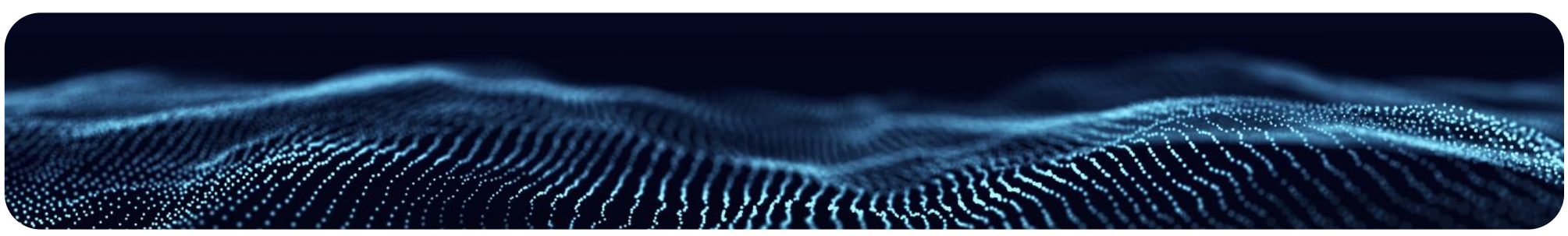
IV. Facteurs de risque pour la sécurité d'un système d'IA

1. Nature des données : Adapter les mesures de sécurité en fonction de la gravité des conséquences d'une faille de sécurité.

2. Maîtrise des données, modèles et outils : Vérifier la fiabilité des sources et des outils utilisés pour éviter les failles de sécurité.

3. Modalités d'accès au système : Limiter la surface d'attaque en contrôlant les modalités d'accès et les informations fournies par le système.

4. Contexte d'utilisation : Adapter les mesures de sécurité en fonction de la criticité du système et de la maturité technique des utilisateurs.



LES FICHES PRATIQUES IA

GARANTIR LA SÉCURITÉ DU DÉVELOPPEMENT D'UN SYSTÈME D'IA

V. Mesures de sécurité à envisager

1. Mesures sur les données d'entraînement :
   - Vérifier la fiabilité, la qualité, et l'intégrité des données d'entraînement.
   - Journaliser et gérer les versions des jeux de données.
   - Utiliser des données fictives ou de synthèse pour les tests.
   - Chiffrer les sauvegardes et les communications.
   - Contrôler les accès aux données et anonymiser ou pseudonymiser les données.

2. Mesures sur le développement du système :
   - Intégrer la protection des données dès la conception du système.
   - Utiliser des outils et bibliothèques vérifiés et conformes aux bonnes pratiques de sécurité.
   - Recourir à un environnement de développement contrôlé et sécurisé.
   - Mettre en œuvre une procédure de développement et d'intégration continus.
   - Documenter la conception et le fonctionnement du système.
   - Conduire des audits de sécurité internes ou externes.

3. Mesures sur le fonctionnement du système :
   - Informer les utilisateurs des limitations du système et fournir des informations pour interpréter les résultats.
   - Prévoir la possibilité de stopper le système en cas de prise de décision automatisée.
   - Contrôler les sorties des systèmes d'IA pour prévenir les contenus problématiques.

4. Mesures transversales :
   - Suivre un plan d'action pour la sécurité.
   - Constituer une équipe pluridisciplinaire pour le développement.
   - Gérer les habilitations, tracer les accès et analyser les traces pour prévenir et détecter les intrusions.

12

LES FICHES PRATIQUES IA

GARANTIR LA SÉCURITÉ DU DÉVELOPPEMENT D'UN SYSTÈME D'IA

VI. Conclusion

Assurer la sécurité du développement des systèmes d'IA est essentiel pour protéger les données personnelles et respecter les obligations du RGPD. Les mesures recommandées par la CNIL permettent de limiter les risques pour les droits et libertés des personnes concernées tout en garantissant la qualité et la fiabilité des systèmes d'IA.

#HASHTAG
CABINET D'AV0CATS

Arnaud Touati

#HASHTAG
CABINET D'AV0CATS

Matthieu Chassagne

pinacle

“ En créant ce guide, notre objectif était de fournir un outil pratique et complet pour les professionnels de l'IA. La structure en fiches permet une consultation rapide tout en offrant une vue d'ensemble des enjeux juridiques.

Nous avons couvert l'intégralité du cycle de vie d'un projet IA, en insistant sur des points cruciaux comme la détermination du régime juridique et la définition de la finalité. Une attention particulière a été portée aux bases légales, notamment l'intérêt légitime, avec des focus sur des cas spécifiques.

L'approche “protection des données dès la conception” reflète notre conviction que la conformité doit être intégrée à toutes les étapes du développement. Nous avons également inclus des aspects techniques souvent négligés, comme l'annotation des données et la sécurité du développement.

Ce guide se veut à la fois rigoureux juridiquement et accessible, visant à promouvoir un développement responsable de l'IA. Des mises à jour régulières sont prévues pour suivre l'évolution rapide du domaine. ”

Synthèse source CNIL :
Les fiches pratiques IA

VOIR LA SOURCE