

IA : Définir une finalité

08 avril 2024

La constitution d'une base de données contenant des données personnelles pour le développement d'un système d'IA est un traitement de données personnelles qui, en application du RGPD, doit poursuivre une finalité (ou objectif) qui soit déterminée, explicite et légitime. La CNIL vous aide à définir la ou les finalités en tenant compte des spécificités du développement de systèmes d'IA.

Le principe

La finalité du traitement est l'objectif poursuivi par l'utilisation des données personnelles. Cet objectif doit être déterminé, c'est-à-dire établi dès la définition du projet. Il doit également être explicite, c'est-à-dire connu et compréhensible. Il doit enfin être légitime, c'est-à-dire compatible avec les missions de l'organisme.

Les données ne doivent pas être traitées ultérieurement de façon incompatible avec cet objectif initial : le principe de finalité limite la manière dont le responsable du traitement peut utiliser ou réutiliser ces données dans le futur.

L'exigence d'une finalité déterminée, explicite et légitime est particulièrement importante, car elle conditionne l'application d'autres principes du RGPD, dont notamment :

- le principe de transparence : l'objectif du traitement doit être porté à la connaissance des personnes concernées afin qu'elles soient en mesure de connaître la raison de la collecte des données les concernant et de comprendre l'utilisation qui en sera faite ;
- le principe de minimisation : les données sélectionnées doivent être adéquates, pertinentes et limitées à ce qui est nécessaire au regard des objectifs pour lesquels elles sont traitées ;
- le principe de limitation des durées de conservation : les données ne peuvent être conservées que pour une durée limitée, définie selon l'objectif pour lesquelles elles ont été collectées.

En savoir plus : [Définir une finalité](#)

Comment définir la finalité du traitement lorsque l'usage opérationnel est identifié dès la phase de développement ?

Ce cas de figure concerne **les systèmes d'IA qui sont développés pour servir à un usage opérationnel précis en phase de déploiement**. Cela exclut les systèmes d'IA qui sont développés sans qu'un usage opérationnel ne soit défini dès la phase de développement (voir section suivante).

Lorsqu'un système d'IA est développé pour un seul usage opérationnel, on considère que la finalité en phase de développement est directement liée à celle poursuivie par le traitement en phase de déploiement. Il en résulte que si la finalité en phase de déploiement est elle-même déterminée, explicite et légitime, la finalité en phase de développement le sera également.

Dans ce cas, ces deux phases consistent tout de même en des traitements distincts dont la conformité aux obligations du RGPD doit être analysée séparément (en termes notamment d'identification de la base légale, d'information des personnes, de minimisation des données collectées, de définition des durées de conservation, etc.).

Exemple : un organisme souhaite constituer une base de données composée de photos de rames de trains en service – c'est-à-dire avec des personnes présentes – afin d'entraîner un algorithme en vue de mesurer l'affluence et la fréquentation des trains à quai dans les gares. La finalité en phase de développement peut être considérée comme déterminée, explicite et légitime au regard de l'usage opérationnel identifié.

Dans certains cas, un système d'IA peut être développé pour plusieurs usages opérationnels définis dès la phase de développement. Dans ce cas, le développement d'un tel système d'IA peut poursuivre plusieurs finalités correspondant aux usages opérationnels identifiés (un traitement de données peut en effet poursuivre simultanément plusieurs finalités si elles sont toutes déterminées, explicites et légitimes).

Comment définir la finalité du traitement pour le développement de systèmes d'IA à usage général ?

Ce cas concerne **les systèmes d'IA dont l'usage opérationnel en phase de déploiement n'est pas clairement identifié dès la phase de développement**. Sont visés, ici, les systèmes d'IA à usage général et les [modèles de fondation](#) utilisables pour une grande variété d'applications pour lesquels il peut être difficile de définir une finalité suffisamment déterminée et explicite au stade du développement.

Exemples :

- Un organisme peut constituer une base de données pour l'entraînement d'un modèle de classification d'images (personnes, véhicules, aliments, etc.) et le rendre publiquement accessible, sans qu'aucun usage opérationnel spécifique ne soit prévu lors du développement du modèle.
- Ce modèle peut être librement réutilisé conformément à la licence associée (en étant éventuellement adapté, par exemple à l'aide de techniques d'[apprentissage par transfert](#) ou *transfer learning*), et à la réglementation relative au droit à l'image et à la propriété intellectuelle, par des organismes tiers pour le développement de systèmes de vision par ordinateur. Les finalités du système d'IA peuvent être variées : détection de personnes par des systèmes de caméras augmentées pour la mesure de l'affluence sur des quais de gare ou encore détection de défauts sur des images prises dans le cadre de contrôles de la qualité de produits.
- Un organisme constitue une base de données pour l'entraînement d'un modèle de langage permettant d'identifier le registre de langue d'un texte. Ce modèle peut être utilisé pour diverses tâches : la rédaction et la relecture d'articles, de courriers, de discours, l'apprentissage du français, etc.

La finalité du traitement en phase de développement peut être considérée comme déterminée, explicite et légitime si elle est suffisamment précise, c'est-à-dire **lorsqu'elle se réfère cumulativement** :

- **au « type » de système développé**, comme, par exemple, le développement d'un modèle de langage de grande taille, d'un système de « vision par ordinateur » ou encore d'un système d'IA générative d'images, de vidéos ou de sons. Les types de systèmes doivent être présentés de manière suffisamment claire et intelligible pour les personnes concernées, compte tenu de leurs complexités techniques et des évolutions rapides dans ce domaine.
- **aux fonctionnalités et capacités techniquement envisageables**, ce qui implique pour le responsable du traitement de dresser une liste des capacités qu'il peut raisonnablement prévoir dès la phase de développement.

Ces critères permettent de prendre en compte le fait que le responsable du traitement ne puisse pas définir au stade du développement d'un système d'IA l'ensemble de ses applications futures, tout en garantissant que le principe de finalité soit respecté.

Exemples de finalités considérées comme explicites et déterminées :

- Développement d'un grand modèle de langage (LLM) capable de répondre à des questions, générer du texte en fonction de contexte (courriels, lettres, rapports, y compris du code informatique), effectuer des traductions, résumés et corrections de texte, faire de la classification de texte, de l'analyse de sentiments, etc. ;
- Développement d'un modèle de reconnaissance vocale capable d'identifier un locuteur, sa langue, son âge, son genre, etc. ;
- Développement d'un modèle de vision par ordinateur capable de détecter différents objets comme des véhicules (voitures, camions, scooters, etc.), des piétons, du mobilier urbain (poubelles, bancs publics, abri-vélos, etc.) ou des éléments de signalisation routière (feux tricolores, panneaux routiers, etc.).

À l'inverse, se référer uniquement au type de système d'IA que l'on souhaite concevoir, sans se référer aux fonctionnalités et capacités techniquement envisageables, ne permet pas de considérer la finalité comme suffisamment précise.

Exemples de finalités qui ne sont pas considérées comme explicites et déterminées :

- Développement d'un modèle d'IA générative (les capacités envisageables ne sont pas définies) ;
- Développement et amélioration d'un système d'IA (ni le type de modèle ni les capacités envisageables ne sont définies) ;
- Développement d'un modèle permettant d'identifier l'âge d'une personne (le « type » n'est pas défini).

Point de vigilance : le responsable du développement du système d'IA à usage général devrait rappeler aux utilisateurs du système leur obligation de définir aussi précisément que possible la finalité pour laquelle le déploiement est prévu et d'en assurer la conformité. Cette conformité dépendra notamment de la prise en compte des risques spécifiques liés à cette finalité. Certains de ces risques devraient être anticipés dès la phase de développement : la CNIL recommande de prendre en compte dès la phase de développement les risques liés aux cas de déploiements connus ou raisonnablement envisageables, quand bien même l'utilisateur du système serait un autre responsable de traitement. Le cas échéant, la licence donnée à des utilisateurs tiers devrait permettre aux personnes concernées de connaître l'étendue de ces risques.

Quelques bonnes pratiques

La transparence du traitement revêt une importance particulière pour les systèmes d'IA à usage général. Ainsi au-delà du respect des obligations rappelées ci-dessus, la CNIL recommande, à titre de bonnes pratiques, que :

- **La finalité mentionne les capacités prévisibles les plus à risque**

Le responsable du traitement est invité à identifier, en amont, les capacités prévisibles du système d'IA qui présentent le plus de risques en phase opérationnelle. Ce serait par exemple le cas des systèmes d'IA identifiés comme « à haut risque » en vertu de la proposition de règlement sur l'intelligence artificielle, actuellement en cours d'adoption.

Ces risques pourront également être pris en compte pour la réalisation de l'AIPD ([voir Fiche n° 5](#)).

- **La finalité mentionne les fonctionnalités exclues par conception**

La description des capacités du système peut inclure les choix de conception du système conduisant à limiter ses fonctionnalités, comme :

- la capacité d'un LLM à ne traiter que du texte court comme des publications sur des réseaux sociaux ;
- le temps de calcul nécessaire à un système de vision par ordinateur pour réaliser ses détections, qui pourrait être trop important pour réaliser de la détection en temps réel ;
- la liste des classes prévues pour un algorithme de classification qui exclurait ainsi la détection d'autres catégories (de sentiments, d'objets, etc.).

Ces limitations pourraient être notamment précisées à l'issue des phases de tests et de validation de la phase de développement, qui permettent au responsable de traitement de préciser le périmètre fonctionnel du système d'IA.

- **La finalité spécifie, dans la mesure du possible, les conditions d'utilisation du système d'IA**

Le responsable du traitement peut spécifier les conditions d'utilisation du système d'IA. Celles-ci peuvent inclure, par exemple, les cas d'usage connus de la solution ou encore les modalités d'utilisation (diffusion du modèle en source ouverte, commercialisation, mise à disposition en SaaS, etc.). Le responsable du traitement pourrait en outre fournir des exemples de cas d'usage opérationnels ou de finalités du système d'IA (par exemple, la régulation du trafic pour un système de vision par ordinateur capable de détecter et quantifier des flux de véhicule).

Comment définir la finalité du développement d'un système d'IA à des fins de recherche scientifique ?

Le responsable du traitement **doit toujours définir l'objectif poursuivi** par la recherche et le traitement de données mis en œuvre. Toutefois, en matière de recherche scientifique, **il peut être admis que le degré de précision de cet objectif soit moins précis ou que les finalités de recherche ne soient pas spécifiées dans leur intégralité**, compte tenu des difficultés que les chercheurs peuvent avoir à la cerner entièrement dès le début de leurs travaux. Il sera alors possible de fournir des informations pour préciser l'objectif à mesure que le projet progresse.

Des dérogations possibles

Les traitements de données effectués à des fins de recherche scientifique bénéficient de dérogations et d'aménagements à certaines obligations (par exemple, s'agissant des [droits des personnes](#), des [durées de conservation](#), ou encore s'agissant du [traitement des données sensibles dans certains cas particuliers](#)).

► Pour en savoir plus : [La recherche scientifique \(hors santé\)](#)

Rappel : qu'est-ce qu'une « recherche scientifique » au sens du RGPD ?

La notion de « recherche scientifique » bénéficie d'une acception large dans le RGPD. En synthèse, la recherche a pour objet de produire des connaissances nouvelles dans tous les domaines dans lesquels la méthode scientifique est applicable.

Afin d'aider les responsables de traitement à déterminer s'ils peuvent bénéficier des dispositions relatives à la recherche scientifique, la CNIL propose un faisceau de critères permettant d'aider le responsable de traitement à déterminer si le traitement qui poursuit une finalité de recherche, relève de la recherche scientifique :

- Dans certains cas, il sera possible de **présumer que la constitution de bases de données d'apprentissage pour l'IA poursuit une finalité de recherche scientifique en raison de la nature de l'organisme** (par exemple, une université ou un centre de recherche public) **ou du mode de financement** (par exemple, financement par l'Agence nationale de la Recherche, ANR).
- À défaut, notamment pour la recherche scientifique privée ne bénéficiant pas de financement public, **il convient d'examiner conjointement les critères suivants** (fondés sur le [Manuel de Frascati de l'OCDE](#) et sur sa définition de la R&D). Ces critères étant cumulatifs, le responsable de traitement devra en principe démontrer qu'ils sont tous remplis pour que le traitement puisse être considéré comme relevant de la recherche scientifique au sens du RGPD. Lorsque ce n'est pas le cas, une analyse au cas par cas est nécessaire pour qualifier le traitement.
 - **La nouveauté** : le traitement doit viser à obtenir des résultats nouveaux (une nouveauté pouvant aussi résulter d'un projet qui amène à constater des divergences potentielles avec le résultat censé être reproduit). L'objet de la recherche peut aider à la qualification de la recherche scientifique. À cet égard, la publication d'articles dans une revue à comité de lecture ou l'octroi d'un brevet permet de qualifier le critère de nouveauté.
 - **La créativité** : ce critère repose sur des notions et hypothèses originales et non évidentes – l'apport des travaux à la connaissance scientifique ou à l'état de la technique. Le développement d'un savoir collectif qui ne profite pas seulement à l'entité morale porteuse du projet de recherche est un indice fort pour qualifier celle-ci de scientifique.
 - **L'incertitude** : le traitement doit revêtir un caractère incertain quant au résultat final.
 - **La systématité** : le traitement doit s'inscrire dans une planification et une budgétisation et mettre en œuvre une méthodologie scientifique. Le respect de normes sectorielles pertinentes de méthodologie et d'éthique est un indice fort pour qualifier la recherche de scientifique. C'est par exemple le cas des exigences méthodologiques particulières pour les [traitements mis en œuvre à des fins de recherche, d'étude ou d'évaluation dans le domaine de la santé](#) qui résultent notamment des articles 72 et suivants la loi « informatique et libertés ».
 - **La transférabilité/reproductibilité** : le traitement doit déboucher sur des résultats qu'il est possible de reproduire ou de transférer dans un champ plus large que celui de la recherche mise en œuvre. À titre d'exemple, la publication de l'étude réalisée et la présentation de la méthodologie de recherche adoptée est un indice fort permettant de souligner la volonté de partage du ou des porteurs de projet.

Exemple :

Pourrait être considéré comme poursuivant des fins de recherche scientifique le développement d'un système d'IA pour une preuve de concept destinée à démontrer la robustesse d'un apprentissage automatique nécessitant moins de données d'entraînement, qui s'inscrirait dans une démarche scientifique documentée ayant vocation à faire l'objet d'une publication.

En savoir plus :

[La réutilisation de données publiquement accessibles à des fins de recherche scientifique \(hors santé\)](#), extrait du guide.

[< Fiche précédente : Déterminer le régime juridique applicable](#)

[Sommaire](#)

[Fiche suivante : Déterminer la qualification juridique des fournisseurs de systèmes d'IA >](#)