

Déterminer la qualification juridique des acteurs

08 avril 2024

Les organismes constituant des bases de données d'apprentissage contenant des données personnelles doivent déterminer leur qualification au sens du RGPD : ils peuvent être responsable de traitement, responsable conjoint ou sous-traitant.

Plusieurs acteurs peuvent intervenir dans le développement d'un système d'IA, avec divers degrés d'implication sur les traitements de données personnelles. Il y a notamment :

- le fournisseur de système d'IA qui développe ou fait développer un système et qui le met sur le marché ou le met en service sous son propre nom ou sa propre marque, à titre onéreux ou gratuit.
- les importateurs, distributeurs, et les utilisateurs de ces systèmes (entendus comme les personnes déployant les systèmes d'IA).

La qualification des acteurs impliqués pour chaque traitement, au sens du RGPD, doit faire l'objet d'une analyse au cas par cas.

Le responsable du traitement

Le principe

Le responsable du traitement est la personne physique ou morale qui détermine les objectifs et les moyens du traitement, c'est-à-dire qui décide du « pourquoi » et du « comment » de l'utilisation de données personnelles.

Les moyens essentiels du traitement sont ceux qui sont étroitement liés à l'objectif et à la portée du traitement, tels que le type de données personnelles qui sont collectées et utilisées, les supports matériels et logiciels utilisés pour le traitement ainsi que leur sécurisation, la durée du traitement, les catégories de destinataires et les catégories de personnes concernées.

En pratique

Certains indices peuvent aider à mener l'analyse au cas par cas pour déterminer qui est responsable du traitement.

Un fournisseur qui est à l'initiative du développement d'un système d'IA et qui en **constitue la base de données** d'apprentissage à partir de données qu'il a sélectionnées pour son propre compte, peut être qualifié de **responsable de traitement**.

Il en ira de même d'un fournisseur qui confie la constitution d'une telle base à un prestataire à travers des instructions documentées suffisamment précises (voir le rôle de sous-traitant ci-dessous).

A noter que dans certains cas, un fournisseur aura recours à un prestataire qui a déjà constitué un jeu de données en tant que responsable du traitement (de sa propre initiative). Il conviendra alors d'identifier les traitements pour lesquels le fournisseur est responsable, comme la réutilisation pour son propre compte d'un jeu de données déjà constitué.

Exemples de responsables de traitement :

- Une plateforme de vidéo à la demande souhaite développer un système d'IA de recommandations. Pour cela, il réutilise une base de données sur ses clients ayant été initialement collectée à des fins de fourniture du service.
La plateforme de vidéo à la demande qui constitue la base de données pour entraîner son système d'IA de recommandations est responsable de ce nouveau traitement puisqu'elle a décidé de l'objectif (entraîner un système d'IA de recommandations) et des moyens essentiels du traitement (à savoir la base de données collectée pour une autre finalité).
- Le fournisseur d'un agent conversationnel qui entraîne son [modèle de langage](#) (« *Large Language Model* » ou LLM en anglais) à partir de données publiquement accessibles sur Internet, est responsable de traitement de la réutilisation des données personnelles publiquement accessibles sur Internet. En effet, il décide à la fois de l'objectif (proposer un agent conversationnel) et des moyens essentiels du traitement (sélectionner les données qu'il va réutiliser).
- Un fournisseur développe un système d'IA sur la base d'un modèle pré-entraîné avec des données personnelles. Il entend le ré-entraîner ou l'ajuster (*fine-tuning* ou *transfer learning*) avec un jeu de données qu'il a lui-même constitué, à son initiative. Dans un tel cas, ce fournisseur devra être qualifié de responsable de traitement, dès lors qu'il poursuit une finalité qui lui est propre et pour laquelle il détermine lui-même les moyens essentiels.

La réutilisation de données collectées par un autre organisme

Lorsque le fournisseur entraîne son système d'IA avec des données collectées par un autre organisme, il est nécessaire de distinguer :

- **le diffuseur des données** : la personne physique ou morale, publique ou privée, qui met à disposition des données personnelles ou une base de données personnelles à des fins de réutilisation ;
- **le réutilisateur des données** : la personne physique ou morale, publique ou privée, traitant ces données ou bases de données en vue d'une exploitation de celles-ci pour son propre compte.

Le diffuseur et le réutilisateur des données sont, en principe, responsables de traitements distincts puisque chacun détermine les objectifs et les moyens essentiels de son propre traitement.

Le diffuseur des données est, en principe, responsable de traitement de la diffusion, alors que le fournisseur du système d'IA qui réutilise les données, est lui responsable de traitement de la réutilisation. Le diffuseur n'est pas, en principe, responsable des traitements de réutilisation de ses données. Il peut toutefois prévoir des conditions d'utilisation des données diffusées pour en limiter les réutilisations ou prévoir certaines dispositions.

Exemple :

Une administration rend publiques et librement réutilisables (open data) des données immobilières. Une entreprise souhaite réutiliser ces données pour constituer une base de données d'apprentissage afin de développer un système d'IA consistant à prédire certaines évolutions immobilières sur un territoire donné. Le diffuseur et le réutilisateur sont alors responsables de traitement distincts, dès lors que ces deux traitements sont indépendants.

En savoir plus : Fiche n° 1 [du guide sur l'ouverture et la réutilisation de données publiquement accessibles](#).

Les responsables conjoints du traitement

Le principe

Lorsque deux responsables du traitement ou plus, déterminent conjointement les finalités et les moyens du traitement, ils sont responsables conjoints du traitement.

Cette qualification peut être plus délicate en présence de plusieurs acteurs exerçant une influence sur la détermination des finalités et des moyens du traitement. Les acteurs doivent notamment déterminer s'ils traitent les données pour des objectifs propres et distincts ou pour un objectif commun.

En pratique

Lorsqu'une base de données d'apprentissage d'un système d'IA est **alimentée par plusieurs responsables de traitement** pour un **objectif conjointement défini**, ces derniers peuvent être qualifiés de **responsables conjoints** du traitement.

Exemples :

Cas n°1 : des centres hospitaliers universitaires développant un système d'IA pour l'analyse de données d'imagerie médicale choisissent de recourir à un même protocole d'[apprentissage fédéré](#). Ce dernier leur permet d'exploiter des données pour lesquelles ils sont initialement des responsables de traitement distincts, afin de ne pas s'en rendre mutuellement destinataires.

Ils déterminent ensemble l'objectif (entraîner un système d'IA d'imagerie médicale) et les moyens de ce traitement (à travers le choix du protocole et la détermination des données qu'ils exploitent) : ils sont donc responsables conjoints de ce traitement d'apprentissage.

Cas n°2 : un consortium composé d'une commune, d'une société fournissant un logiciel de traitement automatisé d'images et une société fournissant des dispositifs vidéo mettent en œuvre une expérimentation visant à installer des caméras augmentées pour enregistrer et analyser le flux et le comportement des véhicules empruntant une voie de circulation au sein de la commune. Le contrat passé entre la ville et les deux sociétés prévoit l'utilisation du logiciel par la commune en conditions

réelles et la possibilité pour les sociétés d'améliorer le logiciel de traitement automatisé d'images par les données collectées en temps réel. Cette amélioration du logiciel de traitement automatisé bénéficie aussi bien à la commune qu'aux sociétés fournissant le logiciel de traitement automatisé d'images et les dispositifs vidéo.

La commune et les deux sociétés seraient ainsi responsables conjoints du traitement de constitution de la base de données d'apprentissage du logiciel de traitement automatisé d'images dès lors qu'ils décident conjointement de la finalité et des moyens essentiels du traitement et que les sociétés n'agissent pas uniquement pour le compte de la commune. En effet, il est possible de considérer qu'ils décident conjointement des moyens essentiels du traitement (en choisissant d'alimenter la base de données d'apprentissage du système d'IA par les données collectées en temps réel par les caméras augmentées et par les données déjà collectées par la société fournissant le logiciel de traitement automatisé d'images) et de l'objectif du traitement (entraîner de manière expérimentale un système d'IA qui permet de détecter les comportements particuliers de véhicules et améliorer le logiciel de traitement automatisé d'images).

A l'inverse, si l'une des sociétés entend réutiliser les données pour une finalité propre dont elle serait la seule à bénéficier (par exemple dans un cadre de recherche et développement), il pourrait alors être envisagé de considérer qu'elle soit responsable d'un traitement distinct.

En cas de responsabilité conjointe, les responsables de traitement conjoints doivent s'assurer de la licéité du traitement (c'est-à-dire de sa conformité à la loi), notamment en définissant de manière transparente leurs obligations respectives dans le cadre d'un accord. La forme de cet accord n'est pas précisée par le RGPD. L'accord doit refléter les rôles de chacune des parties prenantes, les responsables conjoints devant préciser de manière claire « qui fait quoi » pour assurer la protection des données traitées.

À noter : indépendamment des termes de l'accord, la personne concernée par le traitement peut exercer ses droits à l'égard de chacun des responsables conjoints du traitement.

Le recours à un sous-traitant

Le principe

Le sous-traitant est la personne physique ou morale qui traite des données pour le compte du responsable de traitement, dans le cadre d'un service ou d'une prestation.

En pratique

La qualification du fournisseur de système d'IA doit être appréciée au cas par cas.

Un fournisseur de système d'IA peut être sous-traitant lorsqu'il développe un système d'IA pour le compte d'un de ses clients dans le cadre d'une prestation. Le client est pour sa part responsable de traitement dès lors qu'il détermine la finalité et les moyens du traitement.

Dans d'autres configurations, le fournisseur de système d'IA peut être responsable de traitement des systèmes qu'il conçoit pour les commercialiser.

Un fournisseur de système d'IA peut faire appel à un prestataire pour collecter et traiter les données selon ses instructions documentées (par exemple, de collecter des données publiquement accessibles sur Internet, de réutiliser une base de données spécifique mise à disposition en ligne, etc.). **Ce dernier est alors qualifié de sous-traitant.** Il est essentiel pour le fournisseur du système d'IA, en tant que responsable du traitement, de s'assurer que son sous-traitant respecte le RGPD et limite le traitement des données à ses instructions, notamment en concluant un [contrat de sous-traitance](#).

Par ailleurs, le fait d'avoir recours à la même base de données pour différents clients, dans le cadre de prestations de services distinctes, est généralement un indice décisif permettant de considérer que le prestataire est responsable d'un traitement distinct, à minima pour la constitution de la base de données.

Exemple :

Un prestataire s'est vu confier la constitution d'une base de données d'apprentissage par un fournisseur de système d'IA qui lui a indiqué précisément comment elle doit être élaborée (en particulier s'agissant des sources et des catégories de données, avec des exigences de qualité et de documentation). Ce prestataire agira vraisemblablement en qualité de sous-traitant.

A l'inverse, un prestataire qui aurait, à son initiative, constitué un jeu de données qu'il exploite en développant des systèmes d'IA adaptés au besoin de chacun de ses clients sera vraisemblablement responsable du traitement de ce jeu de données, indépendamment de son rôle dans les traitements spécifiques réalisés pour ces clients (qu'il pourrait mettre en œuvre en tant que sous-traitant, par exemple sur la base de données fournies par les clients eux-mêmes).

[< Fiche précédente : Définir une finalité](#)

[Sommaire](#)

[Fiche suivante : Assurer que le traitement est licite >](#)