Opinion of the Board (Art. 64)

Avis 28/2024 sur certains aspects de la protection des données liés au traitement des données personnelles dans le contexte de l'IA modèles

Adopté le 17 décembre 2024

Résumé exécutif

Les technologies de l'IA créent de nombreuses opportunités et avantages dans un large éventail de secteurs et de réseaux sociaux. activités.

En protégeant le droit fondamental à la protection des données, le RGPD soutient ces opportunités et promeut d'autres droits fondamentaux de l'UE, notamment le droit à la liberté de pensée, d'expression et d'information, le droit à l'éducation ou la liberté d'entreprise. Le RGPD constitue ainsi un cadre juridique qui encourage l'innovation responsable.

Dans ce contexte, compte tenu des questions de protection des données soulevées par ces technologies, l'autorité de contrôle irlandaise a demandé au CEPD d'émettre un avis sur des questions d'application générale conformément à l'article 64(2) du RGPD. La demande concerne le traitement des données personnelles dans le cadre des phases de développement et de déploiement de modèles d'intelligence artificielle (« IA ») . Plus en détail, la demande demandait : (1) quand et comment un modèle d'IA peut être considéré comme « anonyme » ; (2) comment les responsables du traitement
peut démontrer la pertinence de l'intérêt légitime comme base juridique dans les phases de développement et (3) de déploiement ; et (4) quelles sont les conséquences du traitement illicite de données à caractère personnel dans la phase de développement d'un modèle d'IA sur le traitement ou l'exploitation ultérieur du modèle d'IA.

En ce qui concerne la première question, l'avis mentionne que les allégations d'anonymat d'un modèle d'IA devraient être évaluées au cas par cas par des autorités de contrôle compétentes, car le CEPD considère que les modèles d'IA formés avec des données à caractère personnel ne peuvent pas, dans tous les cas, être considérés comme anonymes. Pour qu'un modèle d'IA soit considéré comme anonyme, à la fois (1) la probabilité d'extraction directe (y compris probabiliste) de données à caractère personnel concernant les personnes dont les données à caractère personnel ont été utilisées pour développer le modèle et (2) la probabilité d'obtenir, intentionnellement ou non, ces données à caractère personnel à partir de requêtes, devraient être insignifiantes, compte tenu de « tous les moyens raisonnablement susceptibles d'être utilisés » par le responsable du traitement ou une autre personne.

Pour mener leur évaluation, les autorités de contrôle doivent examiner la documentation fournie par le responsable du traitement pour démontrer l'anonymat du modèle. À cet égard, l'avis fournit une liste non prescriptive et non exhaustive de méthodes qui peuvent être utilisées par les responsables du traitement pour démontrer l'anonymat, et donc être prises en compte par les autorités de contrôle lors de l'évaluation de la revendication d'anonymat d'un responsable du traitement. Cela couvre, par exemple, les approches adoptées par les responsables du traitement, pendant la phase de développement, pour empêcher ou limiter la collecte de données personnelles utilisées pour la formation, pour réduire leur identifiabilité, pour empêcher leur extraction ou pour fournir une assurance concernant la résistance aux attaques de l'état de l'art.

En ce qui concerne les deuxième et troisième questions, l'avis fournit des considérations générales que les autorités de contrôle doivent prendre en compte pour évaluer si les responsables du traitement peuvent s'appuyer sur un intérêt légitime en tant que motif. base juridique appropriée pour le traitement effectué dans le cadre du développement et du déploiement de modèles d'IA.

L'avis rappelle qu'il n'existe pas de hiérarchie entre les bases juridiques prévues par le RGPD et qu'il appartient aux responsables du traitement d'identifier la base juridique appropriée pour leurs activités de traitement. L'avis rappelle ensuite le test en trois étapes qui doit être effectué lors de l'évaluation du recours à l'intérêt légitime comme base juridique, à savoir (1) identifier l'intérêt légitime poursuivi par le responsable du traitement ou un tiers ; (2) analyser la nécessité du traitement aux fins de l'intérêt(s) légitime(s) poursuivi(s) (également appelé « test de nécessité ») ; et (3) évaluer que l'intérêt(s) légitime(s) n'est(ne sont) pas supplanté(s) par les intérêts ou les droits et libertés fondamentaux des personnes concernées (également appelé « test de mise en balance »).

En ce qui concerne la première étape, l'avis rappelle qu'un intérêt peut être considéré comme légitime si les trois critères cumulatifs suivants sont remplis : l'intérêt (1) est licite ; (2) est clairement et précisément articulé ; et (3) est réel et actuel (c'est-à-dire non spéculatif). Un tel intérêt peut porter, par exemple, dans le cadre du développement d'un modèle d'IA - développer le service d'un agent conversationnel pour assister les utilisateurs, ou dans son déploiement - améliorer la détection des menaces dans un système d'information.

En ce qui concerne la deuxième étape, l'avis rappelle que l'évaluation de la nécessité implique
en considérant : (1) si l'activité de traitement permettra de poursuivre l'intérêt légitime ; et (2) s'il n'existe pas de moyen moins intrusif de poursuivre cet intérêt. Lorsqu'elles évaluent si la condition de nécessité est remplie, les autorités de contrôle doivent accorder une attention particulière à la quantité de données à caractère personnel traitées et à la proportionnalité de ce traitement par rapport à l'intérêt légitime en jeu, également à la lumière du principe de minimisation des données.

En ce qui concerne la troisième étape, l'avis rappelle que le test de mise en balance doit être effectué en tenant compte des circonstances spécifiques de chaque cas. Il donne ensuite un aperçu des éléments que les autorités de contrôle peuvent prendre en compte pour évaluer si l'intérêt d'un responsable du traitement ou d'un tiers est en jeu.
au-delà des intérêts, des droits fondamentaux et des libertés des personnes concernées.

Dans le cadre de la troisième étape, l'avis met en évidence les risques spécifiques pour les droits fondamentaux qui peuvent apparaître lors des phases de développement ou de déploiement des modèles d'IA. Il précise également que le traitement des données à caractère personnel qui a lieu pendant les phases de développement et de déploiement des modèles d'IA peut avoir des répercussions différentes sur les personnes concernées, qui peuvent être positives ou négatives. Pour évaluer ces répercussions, les autorités de contrôle peuvent tenir compte de la nature des données traitées par les modèles, du contexte du traitement et des éventuelles conséquences ultérieures du traitement.

L'avis souligne également le rôle des attentes raisonnables des personnes concernées dans le test de mise en balance. Cela peut être important en raison de la complexité des technologies utilisées dans les modèles d'IA et du fait qu'il peut être difficile pour les personnes concernées de comprendre la diversité de leurs utilisations potentielles, ainsi que les différentes activités de traitement impliquées. À cet égard, tant les informations fournies aux personnes concernées et le contexte du traitement peuvent faire partie des éléments à prendre en compte pour évaluer si les personnes concernées peuvent raisonnablement s'attendre à ce que leurs données personnelles soient traitées. En ce qui concerne le contexte, cela peut inclure : si les données personnelles étaient ou non accessibles au public, la nature de la relation entre la personne concernée et le responsable du traitement (et s'il existe un lien entre les deux),
la nature du service, le contexte dans lequel les données personnelles ont été collectées, la source à partir de laquelle les données ont été collectées (c'est-à-dire le site Web ou le service où les données personnelles ont été collectées et les paramètres de confidentialité qu'ils offrent), les utilisations potentielles ultérieures du modèle et si les personnes concernées sont réellement conscientes que leurs données personnelles sont en ligne.

L'avis rappelle également que, lorsque les intérêts, les droits et les libertés des personnes concernées semblent prévaloir sur les intérêts légitimes poursuivis par le responsable du traitement ou un tiers, le responsable du traitement peut envisager d'introduire des mesures d'atténuation pour limiter l'impact du traitement sur ces personnes concernées.
Les mesures d'atténuation ne doivent pas être confondues avec les mesures que le responsable du traitement est légalement tenu d'adopter de toute façon pour assurer le respect du RGPD. En outre, les mesures doivent être adaptées aux circonstances de l'espèce et aux caractéristiques du modèle d'IA, y compris à son utilisation prévue. À cet égard, l'avis fournit une liste non exhaustive d'exemples de mesures d'atténuation en relation avec la phase de développement (également en ce qui concerne le web scraping) et la phase de déploiement. Les mesures d'atténuation peuvent être soumises à une évolution rapide et doivent être adaptées aux circonstances de l'espèce.
Il appartient donc aux AS d'évaluer au cas par cas la pertinence des mesures d'atténuation mises en œuvre.

En ce qui concerne la quatrième question, l'avis rappelle de manière générale que les autorités de surveillance disposent de pouvoirs discrétionnaires pour évaluer les éventuelles violations et choisir les mesures appropriées, nécessaires et proportionnées, en tenant compte des circonstances de chaque cas individuel. L'avis envisage ensuite trois scénarios.

Dans le scénario 1, les données personnelles sont conservées dans le modèle d'IA (ce qui signifie que le modèle ne peut pas être considéré comme anonyme, comme détaillé dans la première question) et sont ensuite traitées par le même responsable du traitement (par exemple dans le cadre du déploiement du modèle). L'avis indique que la question de savoir si les phases de développement et de déploiement impliquent des finalités distinctes (constituant ainsi des activités de traitement distinctes) et la mesure dans laquelle l'absence de base juridique pour l'activité de traitement initiale affecte la licéité du traitement ultérieur doivent être évaluées au cas par cas, en fonction du contexte de l'affaire.

Dans le scénario 2, les données personnelles sont conservées dans le modèle et sont traitées par un autre responsable du traitement dans le cadre du déploiement du modèle. À cet égard, l'avis indique que les autorités de contrôle devraient tenir compte du fait que le responsable du traitement qui déploie le modèle a procédé à une évaluation appropriée, dans le cadre de ses obligations de responsabilité visant à démontrer la conformité avec l'article 5(1)(a) et l'article 6 du RGPD, pour s'assurer que le modèle d'IA n'a pas été développé en traitant illégalement des données personnelles. Cette évaluation devrait tenir compte, par exemple, de la source des données personnelles et de la question de savoir si le traitement au cours de la phase de développement a fait l'objet d'une constatation d'infraction, en particulier si elle a été déterminée.
par une AS ou un tribunal, et devraient être plus ou moins détaillées en fonction des risques soulevés par le traitement en phase de déploiement.

Dans le scénario 3, un responsable du traitement traite illégalement des données à caractère personnel pour développer le modèle d'IA, puis s'assure qu'elles sont anonymisées, avant que le même responsable du traitement ou un autre responsable du traitement n'initie un autre traitement de données à caractère personnel dans le cadre du déploiement. À cet égard, l'avis indique que s'il peut être démontré que l'exploitation ultérieure du modèle d'IA n'implique pas le traitement de données à caractère personnel, le CEPD considère que le RGPD ne s'appliquerait pas. Par conséquent, l'illicéité du traitement initial ne devrait pas avoir d'impact sur l'exploitation ultérieure du modèle. En outre, le CEPD considère que, lorsque les responsables du traitement traitent ultérieurement des données à caractère personnel collectées pendant la phase de déploiement, après que le modèle a été anonymisé, le RGPD s'appliquerait à ces opérations de traitement. Dans ces cas, l'avis considère que, en ce qui concerne le RGPD, la licéité du traitement effectué dans la phase de déploiement ne devrait pas être affectée par l'illicéité du traitement initial.

Table des matières

Le Comité européen de la protection des données

Vu l'article 63 et l'article 64(2) du règlement 2016/679/UE du Parlement européen et du Conseil,
Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (ci-après « RGPD »),

Vu l'accord EEE et notamment son annexe XI et son protocole 37, tel que modifié par la décision du Comité mixte de l'EEE n° 154/2018 du 6 juillet 20181 ,

Vu les articles 10 et 22 de son règlement intérieur,

Alors que:

(1) Le rôle principal du Comité européen de la protection des données (ci-après le « Comité » ou le « CEPD ») est de garantir l'application cohérente du RGPD dans l'ensemble de l'Espace économique européen (« EEE »).
L'article 64(2) du RGPD prévoit que toute autorité de contrôle , le président du comité ou la Commission peut demander que toute question d'application générale ou produisant des effets dans plusieurs États membres de l'EEE soit examinée par le comité en vue d'obtenir un avis. L'objectif de cet avis est d'examiner une question d'application générale ou produisant des effets dans plusieurs États membres de l'EEE.

(2) L'avis du comité est adopté conformément à l'article 64, paragraphe 3, du RGPD, en liaison avec l'article 10, paragraphe 2, du règlement intérieur du CEPD, dans un délai de huit semaines à compter de la décision du président et de l'autorité de contrôle compétente selon laquelle le dossier est complet. Sur décision du président, ce délai peut être prolongé de six semaines supplémentaires en fonction de la complexité du sujet.

A ADOPTÉ L'AVIS SUIVANT

1 Introduction

1.1 Résumé des faits

1. Le 4 septembre 2024, l'autorité de contrôle irlandaise (l' « AS irlandaise » ou l'« AS requérante ») a demandé au CEPD d'émettre un avis conformément à l'article 64(2) du RGPD concernant les modèles d'IA et le traitement des données à caractère personnel (« la demande »).

2. Le président du conseil d'administration et l'IE SA ont considéré le dossier comme complet le 13 septembre 2024. Le jour ouvrable suivant, le 16 septembre 2024, le dossier a été diffusé par le secrétariat du CEPD. Le président du Conseil, compte tenu de la complexité de l'affaire, a décidé de prolonger le délai légal, conformément à l'article 64(3) du RGPD et à l'article 10(4) du règlement intérieur du CEPD.

3. La demande porte sur certains éléments de la formation, de la mise à jour, du développement et de l'exploitation des modèles d'IA lorsque les données personnelles font partie de l'ensemble de données pertinent. L'IE SA souligne que la demande

[1] Les références aux « États membres » faites tout au long du présent avis doivent être comprises comme des références aux « États membres de l'EEE ». Les références à l'« Union » faites tout au long du présent avis doivent être comprises comme des références à l'« EEE ».

concerne des questions clés qui ont un impact important sur les personnes concernées et les responsables du traitement dans l'EEE, et qu'il n'existe pas à ce stade de position harmonisée parmi les autorités de contrôle nationales2 . La terminologie qui sera utilisée aux fins du présent avis est fournie dans les sections 2.2 et 2.3 ci-dessous.

4. Les questions suivantes ont été posées par l'IE SA :

Question 1 : Le modèle d'IA final, qui a été formé à l'aide de données personnelles, est-il, dans tous les cas, considéré comme ne répondant pas à la définition des données personnelles (telle qu'énoncée à l'article 4(1) du RGPD) ?

Si la réponse à la question 1 est « oui » :

i. À quelle étape des opérations de traitement conduisant à un modèle d'IA les données personnelles ne sont-elles plus traité?

a) Comment peut-on démontrer que le modèle d'IA ne traite pas de données personnelles ?

ii. Existe-t-il des facteurs qui pourraient empêcher le fonctionnement du modèle d'IA final ? considéré comme anonyme ?

un) Si tel est le cas, comment les mesures prises pour atténuer, prévenir ou se protéger contre ces facteurs (afin de garantir que le modèle d'IA ne traite pas de données personnelles) peuvent-elles être démontrées ?

Si la réponse à la question 1 est « non » :

i. Quelles sont les circonstances dans lesquelles cela pourrait se produire ?

un) Si tel est le cas, comment peut-on démontrer les mesures qui ont été prises pour garantir que le modèle d'IA ne traite pas de données personnelles ?

Question 2 : Lorsqu'un responsable du traitement des données s'appuie sur des intérêts légitimes comme base juridique du traitement des données personnelles pour créer, mettre à jour et/ou développer un modèle d'IA, comment ce responsable du traitement doit-il démontrer la pertinence des intérêts légitimes comme base juridique, tant en ce qui concerne le traitement des données de tiers que de première partie ?

i. Quelles considérations ce responsable du traitement devrait-il prendre en compte pour garantir que les intérêts de les personnes concernées dont les données à caractère personnel sont traitées sont correctement mises en balance avec les intérêts du responsable du traitement dans le cadre :

a) Données de tiers

b) Données de première partie

Question 3 : Après la formation, lorsqu'un responsable du traitement des données s'appuie sur des intérêts légitimes comme base juridique pour le traitement des données personnelles effectué dans le cadre d'un modèle d'IA ou d'un système d'IA dont un modèle d'IA fait partie, comment un responsable du traitement doit-il démontrer la pertinence des intérêts légitimes comme base juridique ?

Question 4 : S'il s'avère qu'un modèle d'IA a été créé, mis à jour ou développé à l'aide de données personnelles traitées illégalement, quel est l'impact de cela, le cas échéant, sur la légalité du traitement ou du fonctionnement continu ou ultérieur du modèle d'IA, que ce soit seul ou dans le cadre d'un modèle d'IA ?
Système, où :

[2] Demande, p.1.

i. Le modèle d'IA, seul ou dans le cadre d'un système d'IA, traite-t-il des données personnelles ?

ii. Ni le modèle d'IA, ni le modèle d'IA en tant qu'élément d'un système d'IA, ne traite de données personnelles ?

1.2 Recevabilité de la demande d'avis au titre de l'article 64(2) du RGPD

5. L'article 64(2) du RGPD prévoit notamment que toute AS peut demander que toute question d'application générale ou produisant des effets dans plus d'un État membre soit examinée par le Comité en vue d'obtenir un avis.

6. L'autorité de contrôle requérante a adressé des questions au CEPD concernant les aspects de protection des données dans le contexte des modèles d'IA. Elle a précisé dans la demande que, bien que de nombreuses organisations utilisent désormais des modèles d'IA, y compris des modèles linguistiques à grande échelle (« LLM »), leur fonctionnement, leur formation et leur utilisation soulèvent « un certain nombre de préoccupations de grande envergure en matière de protection des données » qui impactent les personnes concernées dans l'ensemble de l'UE/EEE »[3][4].

7. La demande soulève, en substance, des questions sur (i) l'application du concept de données à caractère personnel ; (ii) le principe de licéité, en ce qui concerne spécifiquement la base juridique de l'intérêt légitime, dans le contexte des modèles d'IA ; ainsi que sur (iii) les conséquences du traitement illicite de données à caractère personnel dans la phase de développement des modèles d'IA, sur le traitement ou le fonctionnement ultérieur du modèle.

8. Le Comité considère que la demande concerne une « question d'application générale » au sens de l'article 64(2) du RGPD. En particulier, la question concerne l'interprétation et l'application de l'article 4(1), de l'article 5(1)(a) et de l'article 6 du RGPD en ce qui concerne le traitement des données à caractère personnel dans le cadre du développement et du déploiement de modèles d'IA. Comme l'a souligné l'autorité de contrôle requérante, l'application de ces dispositions aux modèles d'IA soulève des questions systémiques, abstraites et nouvelles[5] . Le développement et le déploiement rapides de modèles d'IA par un nombre croissant d'organisations soulèvent des questions spécifiques et, comme le souligne la demande, « le CEPD tirera un grand profit de la conclusion d'une position commune sur les questions soulevées par la présente demande, ces questions étant au cœur des travaux prévus du CEPD à court et moyen terme ».[6]. En outre, les technologies de l'IA créent de nombreuses opportunités et avantages dans un large éventail de secteurs et d'activités sociales. En outre, le RGPD constitue un cadre juridique qui encourage l'innovation responsable. Il s'ensuit qu'il existe un intérêt général à procéder à cette évaluation sous la forme d'un avis du CEPD, afin de garantir l'application cohérente de certaines dispositions du RGPD dans le contexte des modèles d'IA.

9. La condition alternative de l'article 64(2) du RGPD fait référence aux éléments « produisant des effets dans plus d'un État membre ». Le CEPD rappelle que le terme « effets » doit être interprété lato sensu et ne se limite donc pas simplement aux effets juridiques[7] . Étant donné que de plus en plus de modèles d'IA sont formés et utilisés par un nombre croissant d'organisations dans l'EEE, ils ont un impact sur un grand nombre de personnes concernées

[3] Demande, p.1.
[4] Ibid.
[5] Demande, p. 2.
[6] Demande, p. 1. Comme mentionné dans le programme de travail du CEPD pour 2024-2025, adopté le 8 octobre 2024, disponible à l'adresse https://www.edpb.europa.eu/system/files/2024-10/edpb work programme 2024-2025 en.pdf, le CEPD prévoit de publier, entre autres, des lignes directrices sur l'anonymisation, la pseudonymisation et le scraping de données dans le contexte de l'IA générative.

[7] CEPD, Document interne 3/2019 sur les orientations internes relatives à l'article 64 (2) du RGPD, adopté le 8 octobre 2019, https://www.edpb.europa.eu/system/files/2022-07/document internaledpb 201903 art64.2 fr.pdf, paragraphe 15, disponible à

dans l'ensemble de l'EEE, dont certains ont déjà fait part de leurs préoccupations à leur AS compétente8 . Par conséquent, le CEPD considère que la question soulevée par l'AS requérante répond également à cette condition.

10. La demande comprend une motivation écrite sur le contexte et les motivations de la soumission des questions au Comité, y compris sur le cadre juridique pertinent. Par conséquent, le Comité considère que la demande est motivée conformément à l'article 10(3) du règlement intérieur du CEPD.

11. Conformément à l'article 64(3) du RGPD9 , Le CEPD ne rend pas d'avis s'il en a déjà rendu un sur la question. Le CEPD n'a pas rendu d'avis sur la même question et n'a pas encore fourni de réponses aux questions découlant de la demande.

12. Pour ces raisons, la Chambre considère que la demande est recevable et que les questions qui en découlent doivent être analysées dans le présent avis (l' « Avis ») adopté en vertu de l'article 64(2) du RGPD.

2 Portée et notions clés

2.1 Portée de l'avis

13. Le Comité convient avec l'AS requérante que, du point de vue de la protection des données, le développement et le déploiement de modèles d'IA soulèvent des questions fondamentales en matière de protection des données. Les questions portent notamment sur : (i) quand et comment un modèle d'IA peut être considéré comme « anonyme » (question 1 de la demande) ; (ii) comment les responsables du traitement peuvent-ils démontrer la pertinence de l'intérêt légitime comme base juridique dans les phases de développement (question 2 de la demande) et de déploiement (question 3 de la demande) ; et (iii) si le traitement illicite de données à caractère personnel dans la phase de développement a des conséquences sur la licéité du traitement ou du fonctionnement ultérieur du modèle d'IA (question 4 de la demande).

14. Le CEPD rappelle que les autorités de contrôle sont chargées de surveiller l'application du RGPD et doivent contribuer à son application cohérente dans toute l'Union10. Il est donc de la compétence des autorités de contrôle d'enquêter sur des modèles d'IA spécifiques et, ce faisant, de procéder à des évaluations au cas par cas.

15. Le présent avis fournit un cadre permettant aux autorités de contrôle compétentes d'évaluer des cas spécifiques dans lesquels (certaines des) questions soulevées dans la demande se poseraient. Le présent avis ne vise pas à être exhaustif, mais plutôt à fournir des considérations générales sur l'interprétation des dispositions pertinentes, dont les autorités de contrôle compétentes devraient tenir le plus grand compte lorsqu'elles utilisent leurs pouvoirs d'enquête. Bien que le présent avis s'adresse aux autorités de contrôle compétentes et porte sur leurs activités et leurs pouvoirs, il est sans préjudice des obligations des responsables du traitement et des sous-traitants en vertu du RGPD. En particulier, conformément au principe de responsabilité consacré à l'article 5(2) du RGPD, les responsables du traitement sont responsables de tous les principes relatifs au traitement des données à caractère personnel qu'ils traitent et sont en mesure de démontrer leur respect.

16. Dans certains cas, des exemples peuvent être fournis dans l'avis, mais compte tenu de la vaste portée des questions incluses dans la demande, ainsi que des différents types de modèles d'IA qui y sont couverts, tous les scénarios possibles ne seront pas pris en compte dans le présent avis. Les technologies associées aux modèles d'IA sont soumises à une évolution rapide ; par conséquent, les considérations du CEPD dans le présent avis doivent être interprétées à la lumière de cette évolution.

[8] Demande, pp. 1-2.
[9] Article 64(3) du RGPD et article 10(4) du règlement intérieur du CEPD.
[10] Article 51(1) du RGPD et article 51(2) du RGPD.

17. Le présent avis n'analyse pas les dispositions ci-dessous, qui peuvent néanmoins jouer un rôle important lorsque évaluer les exigences en matière de protection des données applicables aux modèles d'IA :

- Traitement des catégories particulières de données : Le CEPD rappelle l'interdiction de l'article 9(1) du RGPD concernant le traitement des catégories particulières de données et les exceptions limitées de l'article 9(2) RGPD11. À cet égard, la Cour de justice de l'Union européenne (« CJUE ») a en outre précisé que « lorsqu'un ensemble de données contenant à la fois des données sensibles et des données non sensibles est [...] collecté en bloc sans qu'il soit possible de séparer les éléments de données les uns des autres au moment de la collecte, le traitement de cet ensemble de données doit être considéré comme interdit, au sens de l'article 9, paragraphe 1, du RGPD, s'il contient au moins un élément de données sensible et qu'aucune des dérogations prévues à l'article 9, paragraphe 2, de ce règlement ne s'applique » a souligné que [12] . En outre, la CJUE a également « aux fins de l'application de l'exception prévue à l'article 9, paragraphe 2, point e), du RGPD, il importe de vérifier si la personne concernée avait l'intention, explicitement et par une action positive claire, de rendre les données à caractère personnel en question accessibles au grand public » 13. Ces considérations devraient être prises en compte lorsque le traitement de données à caractère personnel dans le contexte de modèles d'IA implique des catégories particulières de données.
- Prise de décision automatisée, y compris le profilage : les opérations de traitement effectuées dans le cadre de modèles d'IA peuvent relever du champ d'application de l'article 22 du RGPD, qui impose des obligations supplémentaires aux responsables du traitement et offre des garanties supplémentaires aux personnes concernées. Le CEPD rappelle, à cet égard, ses lignes directrices sur la prise de décision individuelle automatisée et le profilage aux fins du règlement 2016/67914.
- Compatibilité des finalités : l'article 6(4) du RGPD prévoit, pour certaines bases juridiques, des critères qu'un responsable du traitement doit prendre en compte pour déterminer si le traitement effectué pour une autre finalité est compatible avec la finalité pour laquelle les données à caractère personnel sont initialement collectées. Cette disposition peut être pertinente dans le contexte du développement et du déploiement de modèles d'IA et de son applicabilité devrait être évalué par les AS.
- Évaluations d'impact relatives à la protection des données (« AIPD ») (articles 35 du RGPD) : les AIPD sont un élément important de la responsabilité, lorsque le traitement dans le cadre de modèles d'IA est susceptible d'entraîner un risque élevé pour les droits et libertés des personnes physiques15.
- Principe de protection des données dès la conception (article 25(1) du RGPD) : la protection des données dès la conception est une garantie essentielle qui doit être évaluée par les AS dans le cadre du développement et du déploiement d'un modèle d'IA.

---

[11] Voir également le rapport du CEPD sur les travaux menés par le groupe de travail ChatGPT, adopté le 23 mai 2024, paragraphe 18 : « En ce qui concerne le traitement de catégories particulières de données à caractère personnel, l'une des exceptions de l'article 9(2) doit en outre être applicable pour que le traitement soit licite. En principe, l'une de ces exceptions peut être l'article 9(2) (e) du RGPD. Toutefois, le simple fait que des données à caractère personnel soient accessibles au public n'implique pas que "la personne concernée a manifestement rendu ces données publiques" [...] ».

[12] Arrêt de la CJUE du 4 juillet 2023, affaire C-252/21, Meta contre Bundeskartellamt (ECLI:EU:C:2023:537), paragraphe 89.

[13] Arrêt de la CJUE du 4 juillet 2023, affaire C-252/21, Meta contre Bundeskartellamt (ECLI:EU:C:2023:537), paragraphe 77.

[14] Lignes directrices du groupe de travail « Article 29 » (« WP29 ») sur la prise de décision individuelle automatisée et le profilage aux fins du règlement 2016/679, telles que révisées en dernier lieu et adoptées le 6 février 2018, approuvées par le CEPD le 25 mai 2018. Voir également l'arrêt de la CJUE du 7 décembre 2023, affaire C-634/21, SCHUFA Holding et autres (ECLI:EU:C:2023:957).

[15] Lignes directrices du WP29 sur l'analyse d'impact relative à la protection des données (AIPD) et la détermination du caractère « susceptible d'engendrer un risque élevé » du traitement aux fins du règlement 2016/679, révisées et adoptées le 4 octobre 2017, approuvées par le CEPD le 25 mai 2018.

2.2 Notions clés

18. À titre de remarque préliminaire, le CEPD souhaite apporter des éclaircissements sur la terminologie et les concepts qu'il utilisations dans le présent avis, et uniquement aux fins du présent avis :

- Les « données de première partie » désignent les données personnelles que le responsable du traitement a collectées à partir des données sujets.
- « Données tierces » désigne les données personnelles que les responsables du traitement n'ont pas obtenues auprès des personnes concernées, mais collectées ou reçues d'un tiers, par exemple auprès d'un courtier en données ou collectées via le scraping Web.
- Le « Web scraping » est une technique couramment utilisée pour collecter des informations à partir de sources en ligne accessibles au public. Les informations récupérées à partir, par exemple, de services tels que des médias d'information, des réseaux sociaux, des forums de discussion et des sites Web personnels peuvent contenir des données personnelles.
- La demande fait référence au « cycle de vie » des modèles d'IA, ainsi qu'à différentes étapes concernant, entre autres, la « création », le « développement », la « formation », la « mise à jour », le « réglage fin », le « fonctionnement » ou le « post-entraînement » des modèles d'IA. Le CEPD reconnaît que, selon les circonstances, de telles étapes peuvent avoir lieu dans le développement et le déploiement de modèles d'IA et peuvent inclure le traitement de données à caractère personnel à diverses fins de traitement. Néanmoins, aux fins du présent avis, le CEPD considère qu'il est important de rationaliser la catégorisation des étapes susceptibles de se produire. Par conséquent,
  aux fins du présent avis, le CEPD fait référence à la « phase de développement » et à la « phase de déploiement ». Le développement d'un modèle d'IA couvre toutes les étapes précédant tout déploiement du modèle d'IA et comprend, entre autres, le développement du code, la collecte des données personnelles de formation, le prétraitement des données personnelles de formation et la formation. Le déploiement d'un modèle d'IA couvre toutes les étapes relatives à l'utilisation d'un modèle d'IA et peut inclure toutes les opérations menées après la phase de développement. Le CEPD reste conscient de la diversité des cas d'utilisation et de leurs conséquences potentielles en termes de traitement des données personnelles ; par conséquent, les autorités de contrôle doivent déterminer si les observations fournies dans le présent avis sont pertinentes pour le traitement qu'elles évaluent.
- Le CEPD souligne également que, lorsque cela est nécessaire, le terme « formation » fait référence à la partie de la phase de développement où les modèles d'IA apprennent à partir des données pour effectuer la tâche prévue (comme expliqué dans la section suivante de cet avis).
- La notion et la portée des modèles d'IA, telles qu'elles sont comprises par le CEPD aux fins du présent avis, est précisé plus en détail dans la section dédiée suivante.

2.3 Modèles d'IA dans le contexte de l'Avis

19. La loi européenne sur l'intelligence artificielle (« loi IA ») [16] définit un « système d'IA » comme « un système basé sur une machine qui est conçu pour fonctionner avec différents niveaux d'autonomie et qui peut faire preuve d'adaptabilité après déploiement, et qui, pour des objectifs explicites ou implicites, déduit, à partir des entrées qu'il reçoit, comment générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels » 17. Le considérant (12) de la loi sur l'IA explique plus en détail la notion de « système d'IA ».
En conséquence, une caractéristique clé des systèmes d'IA est leur capacité à déduire. Les techniques qui permettent

[16] Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées sur l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2014, (UE) n° 168/2015, (UE) n° 168/2016, (UE) n° 168/2017, (UE) n° 168/2018, (UE) n° 168/2019, (UE) n° 168/2021, (UE) n° 168/2022, (UE) n° 168/2023, (UE) n° 168/2024, (UE) n° 168/2025, (UE) n° 168/2026, N° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (loi sur l'intelligence artificielle).).

[17] Article 3(1) de la loi sur l'IA.

Les inférences lors de la construction d'un système d'IA incluent l'apprentissage automatique, les approches basées sur la logique et les connaissances.

20. Les « modèles d'IA », en revanche, ne sont définis qu'indirectement dans la loi sur l'IA : « Bien que les modèles d'IA soient des composants essentiels des systèmes d'IA, ils ne constituent pas des systèmes d'IA à eux seuls. Les modèles d'IA nécessitent l'ajout d'autres composants, comme par exemple une interface utilisateur, pour devenir des systèmes d'IA. Les modèles d'IA sont généralement intégrés dans les systèmes d'IA et en font partie » 18.

21. Le CEPD comprend que la définition d'un modèle d'IA proposée dans la demande est plus étroite que celle de la loi sur l'IA, car elle fait référence au « modèle d'IA » comme « englobant le produit résultant des mécanismes de formation appliqués à un ensemble de données de formation, dans le contexte de l'intelligence artificielle, de l'apprentissage automatique, de l'apprentissage profond ou d'autres contextes de traitement connexes » et précise en outre que « le terme s'applique aux modèles d'IA qui sont destinés à subir une formation, un réglage fin et/ou un développement supplémentaires, ainsi qu'aux modèles d'IA qui ne le sont pas. » [19]

22. Sur cette base, le CEPD a adopté le présent avis en partant du principe qu'un système d'IA s'appuiera sur un modèle d'IA pour atteindre son objectif prévu en incorporant le modèle dans un cadre plus large (par exemple un système d'IA pour le service client pourrait utiliser un modèle d'IA formé sur des données de conversation historiques pour fournir des réponses aux requêtes des utilisateurs).

23. En outre, les modèles d'IA (ou « modèles ») pertinents pour le présent avis sont ceux développés dans le cadre d'un processus de formation. Ce processus de formation fait partie de la phase de développement, au cours de laquelle les modèles apprennent à partir des données pour effectuer la tâche prévue. Par conséquent, le processus de formation nécessite un ensemble de données à partir duquel le modèle identifiera et « apprendra » des modèles. Dans ces cas, le modèle utilisera différentes techniques pour construire une représentation des connaissances extraites de l'ensemble de données de formation. C'est notamment le cas de l'apprentissage automatique.

24. En pratique, tout modèle d'IA est un algorithme dont le fonctionnement est déterminé par un ensemble d'éléments. Par exemple, les modèles d'apprentissage profond se présentent souvent sous la forme d'un réseau neuronal à plusieurs couches constitué de nœuds reliés par des arêtes ayant des poids, qui sont ajustés pendant l'entraînement pour apprendre les relations entre les entrées et les sorties. Les caractéristiques d'un modèle d'apprentissage profond simple
Les paramètres de l'algorithme de classification d'images sont : (i) le type et la taille de chaque couche, (ii) le poids attribué à chaque arête (parfois appelé « paramètres »), (iii) les fonctions d'activation[20] entre les couches, et éventuellement (iv) d'autres opérations qui peuvent se produire entre les couches. Par exemple, lors de l'entraînement d'un modèle simple d'apprentissage profond pour la classification d'images, les entrées (les « pixels de l'image ») seront associées aux sorties, et les poids pourront être ajustés, de manière à produire la bonne sortie la plupart du temps.

25. D'autres exemples de modèles d'apprentissage profond incluent les LLM et l'IA générative, qui sont utilisés par exemple générer du contenu de type humain et créer de nouvelles données.

26. Sur la base des considérations ci-dessus, conformément à la demande, la portée du présent avis ne couvre que le sous-ensemble de modèles d'IA qui sont le résultat d'un apprentissage de ces modèles avec des données personnelles.

3 Sur le bien-fondé de la demande

3.1 Sur la nature des modèles d'IA par rapport à la définition des données personnelles

[18] Considérant 97 de la loi sur l'IA.

[19] Demande, p. 3.

[20] Il s'agit de fonctions qui calculent, en fonction des entrées et des poids, la sortie d'un nœud neuronal qui sera ensuite envoyée à la couche suivante du réseau neuronal.

27. L'article 4(1) du RGPD définit les données à caractère personnel comme « toute information se rapportant à une personne physique identifiée ou identifiable » (c'est-à-dire la personne concernée). En outre, le considérant 26 du RGPD prévoit que les principes de protection des données ne devraient pas s'appliquer aux informations anonymes, à savoir les informations qui ne se rapportent pas à une personne physique identifiée ou identifiable, compte tenu de « tous les moyens raisonnablement susceptibles d'être utilisés » par le responsable du traitement ou une autre personne. Cela comprend : (i) les données qui n'ont jamais été liées à une personne physique identifiée ou identifiable ; et (ii) les données à caractère personnel qui ont été rendues anonymes de telle manière que la personne concernée n'est pas ou plus identifiable.

28. En conséquence, la question 121 de la demande peut être répondue en analysant si un modèle d'IA résultant d'une formation impliquant le traitement de données personnelles doit, dans tous les cas, être considéré comme anonyme.
Sur la base de la formulation de la question, le CEPD fera référence dans cette section au processus de « formation » d'un modèle d'IA.

29. Tout d'abord, le CEPD souhaite formuler les considérations générales suivantes. Les modèles d'IA, qu'ils soient ou non formés avec des données à caractère personnel, sont généralement conçus pour faire des prédictions ou tirer des conclusions, c'est-à-dire qu'ils sont conçus pour inférer. En outre, les modèles d'IA formés avec des données à caractère personnel sont souvent conçus pour faire des inférences sur des individus différents de ceux dont les données à caractère personnel ont été utilisées pour former le modèle d'IA. Cependant, certains modèles d'IA sont spécifiquement conçus pour fournir des données à caractère personnel concernant des individus dont les données à caractère personnel ont été utilisées pour former le modèle, ou pour rendre ces données accessibles d'une manière ou d'une autre. Dans ces cas, ces modèles d'IA incluront intrinsèquement (et généralement nécessairement) des informations relatives à une personne physique identifiée ou identifiable, et impliqueront donc le traitement de données à caractère personnel. Par conséquent, ces types de modèles d'IA ne peuvent pas être considérés comme anonymes. Ce serait le cas, par exemple, (i) d'un modèle génératif affiné sur les enregistrements vocaux d'un individu pour imiter sa voix ; ou (ii) de tout modèle conçu pour répondre avec des données à caractère personnel issues de la formation lorsqu'il est invité à fournir des informations concernant une personne spécifique.

30. Sur la base des considérations ci-dessus, en répondant à la question 1 de la demande, le CEPD se concentre sur la situation des modèles d'IA qui ne sont pas conçus pour fournir des données personnelles liées aux données de formation.

31. Le CEPD considère que, même lorsqu'un modèle d'IA n'a pas été intentionnellement conçu pour produire des informations relatives à une personne physique identifiée ou identifiable à partir des données d'apprentissage, les informations de l'ensemble de données d'apprentissage, y compris les données à caractère personnel, peuvent néanmoins rester « absorbées » dans les paramètres du modèle, à savoir représentées par des objets mathématiques. Elles peuvent différer des points de données d'apprentissage d'origine, mais peuvent néanmoins conserver les informations d'origine de ces données, qui peuvent finalement être extractibles ou obtenues d'une autre manière, directement ou indirectement, à partir du modèle. Chaque fois que des informations relatives à des personnes identifiées ou identifiables dont les données à caractère personnel ont été utilisées pour entraîner le modèle peuvent être obtenues à partir d'un modèle d'IA par des moyens raisonnablement susceptibles d'être utilisés, on peut conclure qu'un tel modèle n'est pas anonyme.

32. À cet égard, la demande indique que « les publications de recherche existantes mettent en évidence un certain potentiel vulnérabilités pouvant exister dans les modèles d'IA et pouvant entraîner le traitement de données personnelles, [22] aussi comme le traitement des données personnelles qui peut avoir lieu lorsque des modèles sont déployés pour être utilisés avec d'autres données, soit via des interfaces de programmation d'applications (« API »), soit via des interfaces « d'invite » 23.

[21] « Le modèle d'IA final, qui a été formé à l'aide de données personnelles, est-il, dans tous les cas, considéré comme ne répondant pas à la définition des données personnelles (telle qu'énoncée à l'article 4(1) du RGPD) ? »

[22] Tels que les attaques par inférence d'adhésion (OWASP) et les attaques par inversion de modèle (OWASP et Veale et al, 2018).

[23] Demande, p.1-2.

33. Dans le même ordre d'idées, les recherches sur l'extraction de données d'entraînement sont particulièrement dynamiques24. Elles montrent qu'il est possible, dans certains cas, d'utiliser des moyens raisonnablement susceptibles d'extraire des données personnelles de certains modèles d'IA, ou simplement d'obtenir accidentellement des données personnelles par le biais d'interactions avec un modèle d'IA (par exemple dans le cadre d'un système d'IA). Les efforts de recherche continus dans ce domaine permettront d'évaluer plus avant les risques résiduels de régurgitation25 et d'extraction de données personnelles dans un cas donné.

34. Sur la base des considérations ci-dessus, le CEPD considère que les modèles d'IA formés à partir de données à caractère personnel ne peuvent pas, dans tous les cas, être considérés comme anonymes. Au contraire, la question de savoir si un modèle d'IA est anonyme doit être évaluée, sur la base de critères spécifiques, au cas par cas.

3.2 Sur les circonstances dans lesquelles les modèles d'IA pourraient être considérés comme anonymes et la démonstration associée

35. Concernant la question 1 de la demande26, le CEPD est invité à clarifier les circonstances dans lesquelles un modèle d'IA, qui a été formé à l'aide de données à caractère personnel, peut être considéré comme anonyme. Concernant la question 1(i) (a) de la demande, le CEPD est invité à clarifier les preuves et/ou documents que les autorités de sécurité doivent prendre en compte pour évaluer si un modèle d'IA est anonyme.[27]

3.2.1 Considérations générales concernant l'anonymisation dans le contexte présent

36. L'utilisation de l'expression « toute information » dans la définition des « données à caractère personnel » au sens de l'article 4(1) du RGPD reflète l'objectif d'attribuer une portée large à ce concept, qui englobe tous les types d'informations à condition qu'elles « se rapportent » à la personne concernée, qui est identifiée ou peut être identifiée directement ou indirectement.

37. Les informations peuvent se rapporter à une personne physique même si elles sont techniquement organisées ou codées (par exemple dans un format lisible uniquement par machine, qu'il soit propriétaire ou ouvert) d'une manière qui ne rend pas immédiatement apparente la relation avec cette personne physique. Dans de tels cas, des applications logicielles peuvent être utilisées pour identifier, reconnaître et extraire facilement des données spécifiques. Cela est particulièrement vrai pour les modèles d'IA où les paramètres représentent des relations statistiques entre les données d'apprentissage, et où il peut être

---

[24] Français Voir, à cet égard, par exemple : (i) Veale M., Binns R., Edwards L., 2018, Algorithms that remember: model inversion attack and data protection law. Phil. Trans. R. Soc. A 376 : 20180083, disponible sur http://dx.doi.org/10.1098/rsta.2018.0083 ; (ii) Brown H., Lee K., Mireshghallah F., Shokri R. et Tramèr F., What Does it Mean for a Language Model to Preserve Privacy?, 2022, ACM Digital Library, FAccT '22, 20 juin 2022, Séoul, République de Corée, disponible sur https://dl.acm.org/doi/abs/10.1145/3531146.3534642 ; (iii) Vassilev A., Oprea A., Fordyce A., Anderson H., Adversarial Machine Learning A Taxonomy and Terminology of Attacks and Mitigations, janvier 2024, National Institute of Standards and Technology, disponible à l'adresse https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-2e2023.pdf ; (iv) Carlini N., Tramèr F., Wallace E., Jagielski M., Herbert-Voss A., Lee K., Roberts A., Brown T., Song D., Erlingsson U., Oprea A., Raffel C., Extraction de données d'apprentissage à partir de grands modèles linguistiques, arXiv:2012.07805v2 [cs.CR] 15 juin 2021, disponible à l'adresse https://arxiv.org/pdf/2012.07805 ; (v) Fredrikson M., Jha S., Ristenpart T., Attaques par inversion de modèle qui exploitent les informations de confiance et contre-mesures de base, ACM Digital Library, 12 octobre 2015, disponible sur https://dl.acm.org/doi/abs/10.1145/2810103.2813677 ; (vi) Zhang Y., Jia R., Pei H., Wang W., Li B., Song D., The Secret Revealer : Attaques par inversion de modèle génératives contre les réseaux neuronaux profonds, arXiv:1911.07135v2 [cs.LG]
18 avril 2020, disponible sur https://arxiv.org/pdf/1911.07135.

[25] Pour un système d'IA basé sur l'IA générative, la régurgitation correspond à la situation où les sorties seraient directement liées aux données de formation.

[26] « Dans quelles circonstances cela pourrait-il se produire ? » « Si tel est

[27] le cas, comment les mesures qui ont été prises pour garantir que le modèle d'IA ne traite pas de données personnelles peuvent-elles être prises ? 'démontré ?'

il est possible d'extraire des données personnelles exactes ou inexactes (car déduites statistiquement), soit directement à partir des relations entre les données incluses dans le modèle, soit en interrogeant ce modèle.

38. Étant donné que les modèles d'IA ne contiennent généralement pas d'enregistrements pouvant être directement isolés ou liés, mais plutôt des paramètres représentant des relations probabilistes entre les données contenues dans le modèle, il peut être possible de déduire28 des informations du modèle, telles que l'inférence d'appartenance, dans des scénarios réalistes.
Par conséquent, pour qu'une AS convienne avec le responsable du traitement qu'un modèle d'IA donné peut être considéré comme anonyme, elle doit au moins vérifier si elle a reçu suffisamment de preuves que, par des moyens raisonnables : (i) les données personnelles, liées aux données de formation, ne peuvent pas être extraites29 du modèle ; et (ii) toute sortie produite lors de l'interrogation du modèle ne concerne pas les personnes concernées dont les données personnelles ont été utilisées pour former le modèle.

39. Trois éléments doivent être pris en compte par les autorités de surveillance pour déterminer si ces conditions sont remplies. accompli.

40. Premièrement, les autorités de surveillance devraient tenir compte des éléments identifiés dans les avis les plus récents du G29 et/ ou dans les lignes directrices du CEPD sur la question. En ce qui concerne l'anonymisation à la date du présent avis, les autorités de surveillance devraient tenir compte des éléments inclus dans l'avis 05/2014 du G29 sur les techniques d'anonymisation (l' « avis 05/2014 du G29 »), qui stipule que s'il n'est pas possible d'isoler, de relier et de déduire des informations à partir de l'ensemble de données supposément anonyme, les données peuvent être considérées comme anonymes30. Il stipule également que « chaque fois qu'une proposition ne répond pas à l'un des critères, une évaluation approfondie des risques d'identification doit être effectuée » 31. Compte tenu de la probabilité d'extraction et d'inférence mentionnée ci-dessus, le C

[28] (i) Carlini N., Chien S., Nasr M., Song S., Terzis A., Tramer F., Attaques par inférence d'appartenance à partir des premiers principes, arXiv:2112.03570, disponible sur https://arxiv.org/abs/2112.03570 ; (ii) Creţu AM, Guépin F., et De Montjoye YA, Attaques d'inférence de corrélation contre les modèles d'apprentissage automatique. eadj9260(2024). Adv.10, DOI:10.1126/sciadv.adj9260 Sci. disponible à https://www.science.org/doi/10.1126/sciadv.adj9260; (iii) Dana L., Pydi MS, Chevaleyre Y., Mémorisation dans les transformateurs à attention unique arXiv:2411.10115v1 [cs.AI] 15 novembre 2024, disponible sur : https://arxiv.org/abs/2411.10115 ; (iv) Gehrke M., Liebenow J., Mohammadi E. & Braun T. et al. Lifting in Support of Privacy-Preserving Probabilistic Inference. Künstl Intell, 13 juin 2024, disponible sur : https://doi.org/10.1007/s13218-024-00851-y ; (v) Hu H., Membership Inference Attacks and Defenses on Machine Learning Models Literature, disponible sur : https://github.com/HongshengHu/membership-inference-machine-learning-literature ; (vi) Nasr M., Carlini N., Hayase J., Jagielski M., Cooper AF, Ippolito D., Choquette-Choo CA, Wallace E., Tramèr F., et Lee K., Scalable Extraction of Training Data from (Production) Language Models, arXiv:2311.17035 28 novembre 2023, disponible sur : https://arxiv.org/abs/2311.17035 ; (vii) Shokri R., Stronati M., Song C., Shmatikov V., Membership Inference Attacks against Machine Learning Models arXiv:1610.05820v2 [cs.CR], 31 mars 2017, disponible sur https://arxiv.org/abs/1610.05820 ; (viii) Staab R., Vero M., Mislav Balunović, Martin Vechev, 2024, Beyond Memorization: Violating Privacy Via Inference with Large Language Models, arXiv:2310.07298v2, 6 mai 2024, disponible sur https://arxiv.org/abs/2310.07298 ; (ix) Wu F., Cui L., Yao S., Yu S., Inference Attacks in Machine Learning as a Service: A Taxonomy, Review, and Promising Directions arXiv:2406.02027v1 [cs.LG], 27 juin 2024, disponible sur https://arxiv.org/abs/2406.02027v1 ; (x) Zhang J., Das D., Kamath G., Tramèr F., Membership Inference Attacks Cannot Prove that a Model Was Training On Your Data arXiv:2409.19798v1, [cs.LG], 29 septembre 2024, disponible sur https://arxiv.org/abs/2409.19798 ; (xi) Zhou Z., Xiang J., Chen C., et Su S., Quantifying and Analyzing Entity-Level Memorization in Large Language Models, arXiv:2308.15727v2 [cs.CL] 5 novembre 2023, disponible sur : https://arxiv.org/abs/2308.15727 .

[29] L'extraction inclut notamment le cas où les données personnelles sont déduites du modèle d'IA lui-même, avec peu d'informations. ou aucune utilisation des interfaces de requête.

[30] Avis WP29 05/2014, p.24.

[31] Avis WP29 05/2014, p.24.

considère que les modèles d'IA sont très susceptibles de nécessiter une évaluation aussi approfondie des risques d'identification.

41. Deuxièmement, cette évaluation devrait être effectuée en tenant compte de « tous les moyens raisonnablement susceptibles d'être utilisés » par le responsable du traitement ou une autre personne pour identifier les individus32, et la détermination de ces moyens devrait être fondée sur des facteurs objectifs, comme expliqué au considérant 26 du RGPD, qui peuvent inclure : a. les caractéristiques des données de formation elles-mêmes, le modèle d'IA et la procédure de formation33 ;
    b. le contexte dans lequel le modèle d'IA est diffusé et/ou traité34 ;
    c. les informations supplémentaires qui permettraient l'identification et qui pourraient être mises à la disposition de la personne donnée;
    d. les coûts et le temps dont la personne aurait besoin pour obtenir ces informations supplémentaires informations (au cas où elles ne leur seraient pas déjà accessibles)35 ; et
    e. la technologie disponible au moment du traitement, ainsi que les avancées technologiques développements36.

42. Troisièmement, les autorités de surveillance devraient déterminer si les responsables du traitement ont évalué le risque d'identification par le responsable du traitement et par différents types d' « autres personnes », y compris des tiers non intentionnels accédant au modèle d'IA, en examinant également s'ils peuvent raisonnablement être considérés comme étant en mesure d'accéder aux données en question ou de les traiter.

43. En résumé, le CEPD considère que, pour qu'un modèle d'IA soit considéré comme anonyme, en utilisant des moyens raisonnables, à la fois (i) la probabilité d'extraction directe (y compris probabiliste) de données personnelles concernant les personnes dont les données personnelles ont été utilisées pour former le modèle ; ainsi que (ii) la probabilité d'obtenir, intentionnellement ou non, de telles données personnelles à partir de requêtes, devraient être insignifiantes37 pour toute personne concernée. Par défaut, les autorités de surveillance devraient considérer que les modèles d'IA sont susceptibles de nécessiter une évaluation approfondie de la probabilité d'identification pour parvenir à une conclusion sur leur éventuelle nature anonyme. Cette probabilité devrait être évaluée en tenant compte de « tous les moyens raisonnablement susceptibles d'être utilisés » par le responsable du traitement ou une autre personne, et devrait également tenir compte de l'utilisation (ré)utilisation ou de la divulgation involontaire du modèle.

3.2.2 Éléments permettant d'évaluer la vraisemblance résiduelle d'identification

44. Bien que des mesures puissent être prises aux stades de développement et de déploiement afin de réduire la probabilité d'obtenir des données personnelles à partir d'un modèle d'IA, l'évaluation de l'anonymat d'un modèle d'IA devrait également prendre en compte l'accès direct au modèle.

45. En outre, les autorités de surveillance devraient évaluer, au cas par cas, si les mesures mises en œuvre par l'
les moyens mis en œuvre par le responsable du traitement pour garantir et prouver qu'un modèle d'IA est anonyme sont appropriés et efficaces.

---

[32] Arrêt de la CJUE du 19 octobre 2016, affaire C-582/14, Breyer contre Bundesrepublik Deutschland (ECLI:EU:C:2016:779), paragraphe 43.

[33] Cela inclut des caractéristiques telles que l'unicité des enregistrements dans les données de formation, la précision des informations, agrégation, randomisation et en particulier comment celles-ci affectent la vulnérabilité à l'identification.

[34] Cela comprend des éléments contextuels, tels que la limitation de l'accès à certaines personnes seulement et des garanties juridiques.

[35] Arrêt de la CJUE du 7 mars 2024, affaire C-479/22 P, OC contre Commission européenne (ECLI:EU:C:2024:215), paragraphe 50.

[36] Arrêt de la CJUE du 7 mars 2024, affaire C-479/22 P, OC contre Commission européenne (ECLI:EU:C:2024:215), paragraphe 50.

[37] Arrêt de la CJUE du 19 octobre 2016, affaire C-582/14, Breyer contre Bundesrepublik Deutschland (ECLI:EU:C:2016:779), paragraphe 46, et arrêt de la CJUE du 7 mars 2024, affaire C-479/22 P, OC contre Commission européenne (ECLI:EU:C:2024:215), paragraph 51.

46. En particulier, la conclusion de l'évaluation d'une AS peut différer entre un modèle d'IA accessible au public, qui est accessible à un nombre inconnu de personnes avec une gamme inconnue de méthodes pour tenter d'extraire des données personnelles, et un modèle d'IA interne accessible uniquement aux employés. Bien que dans les deux cas, les AS doivent vérifier que les responsables du traitement ont rempli leur obligation de responsabilité en vertu de l'article 5(2) et de l'article 24 du RGPD, les « moyens raisonnablement susceptibles d'être utilisés » par d'autres personnes peuvent avoir une incidence sur la portée et la nature des scénarios possibles à prendre en compte. Par conséquent, en fonction du contexte de développement et de déploiement du modèle, les AS peuvent envisager différents niveaux de test et de résistance aux attaques.

47. À cet égard, le CEPD fournit ci-dessous une liste non prescriptive et non exhaustive d'éléments possibles qui peuvent être pris en compte par les autorités de contrôle lors de l'évaluation de la demande d'anonymat d'un responsable du traitement. D'autres approches peuvent être possibles si elles offrent un niveau de protection équivalent, notamment en tenant compte de l'état de la technique.

48. La présence ou l'absence des éléments énumérés ci-dessous ne constitue pas un critère concluant pour évaluer l'anonymat d'un modèle d'IA.

3.2.2.1 Conception du modèle d'IA

49. En ce qui concerne la conception des modèles d'IA, les autorités de contrôle devraient évaluer les approches adoptées par les contrôleurs pendant la phase de développement. L'application et l'efficacité de quatre domaines clés (identifiés ci-dessous) devraient être prises en compte à cet égard.

Sélection des sources

50. Le premier domaine d'évaluation consiste à examiner la sélection des sources utilisées pour former le modèle d'IA. Cela comprend une évaluation, par les AS, de toutes les mesures prises pour éviter ou limiter la collecte de données personnelles, y compris, entre autres, (i) la pertinence des critères de sélection; (ii) la pertinence et l'adéquation des sources choisies compte tenu des objectifs visés; et (iii) si des sources inappropriées ont été exclues.

Préparation et minimisation des données

51. Le deuxième domaine d'évaluation concerne la préparation des données pour la phase de formation. Les autorités de sécurité doivent examiner en particulier : (i) si l'utilisation de données anonymes et/ou personnelles ayant subi une pseudonymisation a été envisagée ; et (ii) lorsqu'il a été décidé de ne pas utiliser de telles mesures, les raisons de cette décision, compte tenu de l'objectif visé ; (iii) les stratégies et techniques de minimisation des données employées pour limiter le volume de données personnelles incluses dans le processus de formation ; et (iv) tout processus de filtrage des données mis en œuvre avant la formation du modèle visant à supprimer les données anonymes et/ou personnelles ayant subi une pseudonymisation.
données personnelles non pertinentes.

Choix méthodologiques concernant la formation

52. Le troisième domaine d'évaluation concerne la sélection de méthodes robustes dans le développement de modèles d'IA. Les SA doivent évaluer les choix méthodologiques susceptibles de réduire considérablement ou d'éliminer l'identifiabilité, notamment : (i) si cette méthodologie utilise des méthodes de régularisation pour améliorer la généralisation du modèle et réduire le surajustement ; et, surtout, (ii) si le contrôleur a mis en œuvre
techniques appropriées et efficaces de préservation de la vie privée (par exemple, confidentialité différentielle).

Mesures concernant les résultats du modèle

53. Le dernier domaine d'évaluation concerne toutes les méthodes ou mesures ajoutées au modèle d'IA lui-même qui peuvent n'a pas d'impact sur le risque d'extraction directe de données personnelles pour le modèle par toute personne y accédant directement, mais cela pourrait réduire la probabilité d'obtenir des données personnelles liées aux données de formation à partir de requêtes.

3.2.2.2 Analyse du modèle d'IA

54. Pour que les autorités de certification puissent évaluer la robustesse du modèle d'IA conçu en matière d'anonymisation, la première étape consiste à s'assurer que la conception a été développée comme prévu et qu'elle est soumise à une gouvernance technique efficace. Les autorités de certification doivent évaluer si les contrôleurs ont effectué des audits basés sur des documents (internes ou externes) qui incluent une évaluation des mesures choisies et de leur impact pour limiter la probabilité d'identification. Cela pourrait inclure l'analyse des rapports de revues de code, ainsi qu'une analyse théorique documentant la pertinence des mesures choisies pour réduire la probabilité de ré-identification du modèle concerné.

3.2.2.3 Test du modèle d'IA et résistance aux attaques

55. Enfin, les autorités de contrôle doivent prendre en considération la portée, la fréquence, la quantité et la qualité des tests que le contrôleur a effectués sur le modèle. En particulier, les autorités de contrôle doivent tenir compte du fait que les tests réussis Les tests portant sur des attaques connues et de pointe ne peuvent être que des preuves de la résistance à ces attaques. À la date de la présente opinion, cela pourrait inclure, entre autres, des tests structurés contre : (i) l'inférence d'attributs et d'appartenance ; (ii) l'exfiltration ; (iii) la régurgitation de données d'entraînement ; (iv) l'inversion de modèle ; ou (v) les attaques de reconstruction.

3.2.2.4 Documentation

56. Les articles 5, 24, 25 et 30 du RGPD et, dans les cas de risque élevé probable pour les droits et libertés des personnes concernées, l'article 35 du RGPD, exigent des responsables du traitement qu'ils documentent de manière adéquate leurs opérations de traitement. Cela s'applique également à tout traitement qui comprendrait la formation d'un modèle d'IA, même si l'objectif du traitement est l'anonymisation. Les autorités de contrôle doivent tenir compte de cette documentation et de toute évaluation régulière des risques qui en découlent pour le traitement effectué par les responsables du traitement, car il s'agit d'étapes fondamentales pour démontrer que les données à caractère personnel ne sont pas traitées.

57. Le CEPD considère que les autorités de surveillance devraient prendre en compte la documentation chaque fois qu'une réclamation l'anonymat concernant un modèle d'IA donné doit être évalué. Le CEPD note que, si une AS n'est pas en mesure de confirmer, après avoir évalué la demande d'anonymat, y compris à la lumière de la documentation, que des mesures efficaces ont été prises pour anonymiser le modèle d'IA, la SA serait en mesure d'envisager que le responsable du traitement n'a pas rempli ses obligations de responsabilité en vertu de l'article 5(2) du RGPD. Par conséquent, le respect des autres dispositions du RGPD doit également être pris en compte.

58. Idéalement, les AS devraient vérifier si la documentation du responsable du traitement comprend :

a. toute information relative aux AIPD, y compris toutes les évaluations et décisions qui ont déterminé qu'une AIPD n'était pas nécessaire ;

b. tout conseil ou commentaire fourni par le délégué à la protection des données (« DPD ») (lorsqu'un DPD était - ou aurait dû être - nommé);

c. des informations sur les mesures techniques et organisationnelles prises lors de la conception du modèle d'IA pour réduire la probabilité d'identification, y compris le modèle de menace et les évaluations des risques sur lesquels ces mesures sont basées. Cela devrait inclure les mesures spécifiques pour chaque source d'ensembles de données de formation, y compris les URL des sources pertinentes et les descriptions des mesures prises (ou déjà prises par des fournisseurs d'ensembles de données tiers) ;

d. les mesures techniques et organisationnelles prises à toutes les étapes du cycle de vie du modèle, qui ont contribué à l'absence de données à caractère personnel dans le modèle ou l'ont vérifiée ;

e. la documentation démontrant la résistance théorique du modèle d'IA aux techniques de réidentification, ainsi que les contrôles conçus pour limiter ou évaluer le succès et l'impact des principales attaques (régurgitation, attaques par inférence d'appartenance, exfiltration, etc.). Cela peut inclure, en

notamment : (i) le rapport entre la quantité de données d'apprentissage et le nombre de paramètres du modèle, y compris l'analyse de son impact sur le modèle38 ; (ii) les mesures de la probabilité de ré-identification sur la base de l'état actuel de la technique ; (iii) les rapports sur la manière dont le modèle a été testé (par qui, quand, comment et dans quelle mesure) et (iv) les résultats des tests ;

f. la documentation fournie au(x) responsable(s) du traitement déployant le modèle et/ou aux personnes concernées, notamment la documentation relative aux mesures prises pour réduire la probabilité d'identification et concernant les éventuels risques résiduels.

3.3 Sur la pertinence de l'intérêt légitime comme base juridique pour le traitement des données personnelles dans le cadre du développement et du déploiement de modèles d'IA

59. Pour répondre aux questions 2 et 3 de la demande, le CEPD fournira d'abord des observations générales sur certains aspects importants que les autorités de contrôle devraient prendre en compte, quelle que soit la base juridique du traitement, lors de l'évaluation de la manière dont les responsables du traitement peuvent démontrer leur conformité avec le RGPD dans le contexte des modèles d'IA. Le CEPD, s'appuyant sur les Lignes directrices 1/2024 sur le traitement des données à caractère personnel fondé sur l'article 6(1)(f) Le RGPD39 examinera ensuite les trois étapes requises par l'évaluation de l'intérêt légitime dans le cadre du développement et du déploiement de modèles d'IA.

3.3.1 Observations générales

60. Le CEPD rappelle que le RGPD n'établit aucune hiérarchie entre les différentes bases juridiques prévues à l'article 6(1) du RGPD40.

61. L'article 5 du RGPD définit les principes relatifs au traitement des données à caractère personnel. Le CEPD met en évidence ceux qui sont importants pour le présent avis et qui devraient au moins être pris en compte par les autorités de surveillance lors de l'évaluation de modèles d'IA spécifiques, ainsi que les exigences les plus pertinentes des autres dispositions du RGPD, compte tenu de la portée du présent avis.

62. Principe de responsabilité (article 5(2) du RGPD) - Ce principe prévoit que le responsable du traitement est responsable du respect du RGPD et doit être en mesure de démontrer ce respect. À cet égard, les rôles et responsabilités des parties qui traitent des données à caractère personnel dans le cadre du développement ou du déploiement d'un modèle d'IA doivent être évalués avant le traitement, afin de définir dès le départ les obligations des responsables du traitement ou des responsables conjoints du traitement, et des sous-traitants (le cas échéant).

63. Principes de licéité, de loyauté et de transparence (article 5(1)(a) du RGPD) - Lors de l'évaluation de la licéité du traitement dans le contexte des modèles d'IA, à la lumière de l'article 6(1) du RGPD, le CEPD estime qu'il est utile de distinguer les différentes étapes du traitement des données à caractère personnel41. Le principe de loyauté, qui est étroitement lié au principe de transparence, exige que les données à caractère personnel ne soient pas traitées par des méthodes déloyales, ou par tromperie, ou d'une manière qui soit « injustifiablement préjudiciable, illicite ou susceptible de nuire à la sécurité des données ».

---

[38] Ricciato F., A Cautionary Reflection on (Pseudo-)Synthetic Data from Deep Learning on Personal Data, conférence Privacy in Statistical Databases (PSD 2024), Antibes, France, septembre 2024, diapositives disponibles à l'adresse suivante : https://cros.ec.europa.eu/system/files/2024-10/20240926 PSD2024 Ricciato v6 1.pdf et Belkin M., Hsu D., Ma S., & Mandal S. (2019), Reconciling modern machine-learning practice and the classic bias–variance trade-off. Proceedings of the National Academy of Sciences, 24 juillet 2019, 116(32) 15849-15854, disponible à l'adresse suivante : https://www.pnas.org/doi/10.1073/pnas.1903070116

[39] Voir les lignes directrices du CEPD 1/2024 sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024.

[40] Ibid, paragraphe 1.

[41] Rapport du CEPD sur les travaux entrepris par le groupe de travail ChatGPT, adopté le 23 mai 2024, paragraphe 14.

42. Compte tenu de la complexité des technologies impliquées, les informations sur le traitement des données à caractère personnel dans les modèles d'IA devraient donc être fournies d'une manière accessible, compréhensible et conviviale43 . La transparence sur le traitement des données à caractère personnel comprend, en particulier, le respect des obligations d'information énoncées aux articles 12 à 14 du RGPD44, qui exigent également, en cas de prise de décision automatisée, y compris le profilage, des informations significatives sur la logique impliquée, ainsi que sur l'importance et les conséquences envisagées du traitement pour la personne concernée45. Sachant que les phases de développement des modèles d'IA peuvent impliquer la collecte de grandes quantités de données à partir de sources accessibles au public (par exemple via des techniques de scraping Web), le recours à l'exception prévue à l'article 14(5)(b) du RGPD est strictement limité au cas où les exigences de cette disposition sont pleinement satisfaites46.

64. Principes de limitation des finalités et de minimisation des données (article 5(1)(b), (c) du RGPD) - Conformément au principe de minimisation des données, le développement et le déploiement de modèles d'IA exigent que les données personnelles soient adéquates, pertinentes et nécessaires au regard de la finalité. Cela peut inclure le traitement de données personnelles pour éviter les risques de biais et d'erreurs potentiels lorsque cela est clairement et spécifiquement identifié dans le cadre de la finalité, et que les données personnelles sont nécessaires à cette finalité (par exemple, elles ne peuvent pas être efficacement atteintes par le traitement d'autres données, y compris des données synthétiques ou anonymisées)47. Le WP29 a déjà souligné que « la finalité de la collecte doit être clairement et spécifiquement identifiée » [...]' 48. Pour évaluer si la finalité poursuivie est légitime, spécifique et explicite, et si le traitement est conforme au principe de minimisation des données, il convient d'abord d'identifier l'activité de traitement en jeu. Notamment, les différentes étapes des phases de développement ou de déploiement peuvent constituer des activités de traitement identiques ou différentes, et peuvent impliquer des responsables successifs ou des responsabilités conjointes contrôleurs. Dans certains cas, il est possible de déterminer la finalité qui sera poursuivie lors du déploiement du modèle d'IA à un stade précoce de développement. Même lorsque ce n'est pas le cas, un certain contexte pour ce déploiement doit déjà être clair et, par conséquent, la manière dont ce contexte informe la finalité du développement doit être prise en compte. Lors de l'examen de la finalité du traitement à un stade donné de développement, les AS doivent s'attendre à un certain degré de détail de la part du ou des contrôleurs et à une explication de la manière dont ces détails informent la finalité du traitement. Cela peut inclure, par exemple, des informations sur le type de modèle d'IA développé, ses fonctionnalités attendues et tout autre contexte pertinent déjà connu à ce stade. Le contexte du déploiement peut également inclure, par exemple, si un modèle est en cours de développement pour un déploiement interne, si le contrôleur a l'intention de le déployer en interne ou non.

---

[42] Rapport du CEPD sur les travaux entrepris par le groupe de travail ChatGPT, adopté le 23 mai 2024, paragraphe 23 ; Lignes directrices du CEPD 4/2019 sur l'article 25 Protection des données dès la conception et par défaut, version 2.0, adoptées le 20 octobre 2020, paragraphe 69 ; Lignes directrices du groupe de travail Article 29 sur la transparence au titre du règlement 2016/679, révisées et adoptées le 11 avril 2018, approuvées par le CEPD le 25 mai 2018, paragraphe 2.

[43] Lignes directrices du groupe de travail Article 29 sur la transparence au titre du règlement 2016/679, révisées et adoptées le 11 Avril 2018, approuvé par le CEPD le 25 mai 2018, paragraphe 5.

[44] Voir également le considérant 39 du RGPD, qui stipule qu'il devrait être « transparent pour les personnes physiques que des données à caractère personnel les concernant sont collectées, utilisées, consultées ou traitées d'une autre manière et dans quelle mesure les données à caractère personnel sont ou seront traitées [...] ».

[45] Article 13(2)(f) du RGPD et article 14(2)(g) du RGPD.

[46] Rapport du CEPD sur les travaux entrepris par le groupe de travail ChatGPT, adopté le 23 mai 2024, paragraphe 27.

[47] En outre, l'article 10(5) de la loi AI prévoit des règles spécifiques pour le traitement de catégories particulières de données à caractère personnel. données relatives aux systèmes d'IA à haut risque afin de garantir la détection et la correction des biais.

[48] Avis 03/2013 du Groupe de travail Article 29 sur la limitation des finalités (WP203), pp. 15-16.

de vendre ou de distribuer le modèle à des tiers après son développement, y compris si le modèle est principalement destiné à être déployé à des fins de recherche ou à des fins commerciales.

65. Droits des personnes concernées (chapitre III du RGPD) - Nonobstant la nécessité pour les AS de garantir que tous les droits des personnes concernées sont respectés lorsque des modèles d'IA sont développés et déployés par les responsables du traitement, le CEPD rappelle que chaque fois qu'un intérêt légitime est invoqué comme base juridique par un responsable du traitement, le droit d'opposition en vertu de l'article 21 du RGPD s'applique et doit être garanti49.

3.3.2 Considérations sur les trois étapes de l'évaluation de l'intérêt légitime dans le contexte du développement et du déploiement de modèles d'IA

66. Afin de déterminer si un traitement donné de données à caractère personnel peut être fondé sur l'article 6(1)(f) Conformément au RGPD, les autorités de contrôle doivent vérifier que les responsables du traitement ont soigneusement évalué et documenté si les trois conditions cumulatives suivantes sont remplies : (i) la poursuite d'un intérêt légitime par le responsable du traitement ou par un tiers ; (ii) le traitement est nécessaire pour poursuivre l'intérêt légitime ; et (iii) l'intérêt légitime n'est pas supplanté par les intérêts ou les droits et libertés fondamentaux des personnes concernées50.

3.3.2.1 Première étape – Poursuite d'un intérêt légitime par le responsable du traitement ou par un tiers

67. Un intérêt est l'enjeu ou l'avantage plus large qu'un responsable du traitement ou un tiers peut avoir en s'engageant dans une activité de traitement spécifique 51. Bien que le RGPD et la CJUE aient reconnu plusieurs intérêts comme étant légitimes52, l'évaluation de la légitimité d'un intérêt donné doit être le résultat d'une analyse au cas par cas.

68. Comme le rappelle le CEPD dans ses Lignes directrices sur l'intérêt légitime53, un intérêt peut être considéré comme légitime si les trois critères cumulatifs suivants sont remplis :

a. L'intérêt est licite54;

---

[49] Conformément à l'article 21 du RGPD, si une personne concernée s'oppose, pour des raisons tenant à sa situation particulière, au traitement des données à caractère personnel la concernant, le responsable du traitement ne doit plus traiter les données à caractère personnel, à moins que le responsable du traitement ne démontre qu'il existe des motifs légitimes et impérieux justifiant le traitement qui prévalent sur les intérêts et les droits et libertés de la personne concernée ou servant à la constatation, à l'exercice ou à la défense de droits en justice. Par conséquent, les deux aspects à prendre en compte par les autorités de contrôle sont de savoir si le responsable du traitement est en mesure de démontrer l'existence de ces motifs légitimes impérieux et impérieux et si le droit d'opposition peut être exercé.

[50] CJUE, arrêt du 4 juillet 2023, affaire C-252/21, Meta c. Bundeskartellamt (ECLI:EU:C:2023:537), point 7. 106 ; CJUE, arrêt du 11 décembre 2019, affaire C-708/18, Asociatia de Proprietari bloc M5A-ScaraA (ECLI:EU:C:2019:1064), paragraphe 40. Voir également les lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0 adoptée le 8 octobre 2024, paragraphes 12 et suivants. Comme le rappellent ces lignes directrices, cette « évaluation doit être effectuée au début du traitement, avec la participation du délégué à la protection des données (DPD) (s'il est désigné), et doit être documentée par le responsable du traitement conformément au principe de responsabilité énoncé à l'article 5(2) du RGPD ».

[51] Lignes directrices du CEPD 1/2024 sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 14.

[52] Lignes directrices du CEPD 1/2024 sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 16.

[53] Lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 17.

[54] CJUE, arrêt du 4 octobre 2024, affaire C-621/22, Royal Dutch Lawn Tennis Association (ECLI:EU:C:2024:857), paragraphe 49, où la CJUE a souligné qu'un intérêt légitime ne peut pas être contraire à la loi. À cet égard, le CEPD souligne que, selon le cas, les cadres législatifs devraient être pris en compte lors de l'évaluation de la licéité d'un intérêt donné. Voir par exemple : l'article 26(3) et l'article 28 du règlement (UE) 2022/2065 du Parlement européen et du Conseil du 19 octobre 2022 relatif à un marché unique

b. L'intérêt est clairement et précisément exprimé; et

c. L'intérêt est réel et présent, et non spéculatif.

69. Sous réserve des deux autres étapes requises par l'évaluation de l'intérêt légitime, les exemples suivants peuvent constituer un intérêt légitime dans le contexte des modèles d'IA : (i) développer le service d'un agent conversationnel pour assister les utilisateurs ; (ii) développer un système d'IA pour détecter les contenus ou comportements frauduleux ; et (iii) améliorer la détection des menaces dans un système d'information.

3.3.2.2 Deuxième étape – Analyse de la nécessité du traitement pour poursuivre l'intérêt légitime

70. La deuxième étape de l'évaluation consiste à déterminer si le traitement des données à caractère personnel est nécessaire aux fins de l'intérêt(s) légitime(s) poursuivi(s)55 (« test de nécessité »).

71. Le considérant 39 du RGPD précise que « les données à caractère personnel ne devraient être traitées que si la finalité du traitement ne peut raisonnablement être atteinte par d'autres moyens ». Selon la CJUE et les orientations précédentes du CEPD, la condition relative à la nécessité du traitement devrait être examinée à la lumière des droits et libertés fondamentaux des personnes concernées, et en conjonction avec le principe de minimisation des données consacré à l'article 5(1)(c) du RGPD56.

72. La méthodologie retenue par la CJUE prend en compte le contexte du traitement ainsi que les effets sur le responsable du traitement et sur les personnes concernées. L'appréciation de la nécessité comporte donc deux éléments : (i) si l'activité de traitement permettra la poursuite de la finalité 57 ; et (ii) s'il n'existe pas de moyen moins intrusif de poursuivre cette finalité 58.

Pour les services numériques et modifiant la directive 2000/31/CE (loi sur les services numériques) (« DSA ») sur la publicité ciblée interdite destinée aux mineurs ; article 5(1) et (2) de la loi sur l'IA sur les pratiques d'IA interdites (pratiques manipulatrices et en dessous du seuil de conscience) ; traitement en violation des droits de propriété intellectuelle et des dispositions de la directive (UE) 2019/790 sur le droit d'auteur et les droits voisins dans le marché unique numérique.

[55] Lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphes 28-30.

[56] CJUE, arrêt du 4 juillet 2023, affaire C-252/21, Meta c. Bundeskartellamt (ECLI:EU:C:2023:537), points 108 et 109, faisant également référence à CJUE, arrêt du 11 décembre 2019, affaire C-708/18, Asociația de Proprietari bloc M5A-ScaraA (ECLI:EU:C:2019:1064), point 48 ; CJUE, arrêt du 9 novembre 2010, affaires jointes C-92/09 et C-93/09, Volker und Markus Schecke (ECLI:EU:C:2010:662), points 85 et 86 ; Français CJUE, arrêt du 22 juin 2021, affaire C-439/19, Latvijas Republikas Saeima (ECLI:EU:C:2021:504), paragraphes 98, 109, 110, 113. Voir aussi par exemple : Lignes directrices du CEPD n° 3/2019 sur le traitement des données à caractère personnel via des appareils vidéo, version 2.0, adoptées le 29 janvier 2020, paragraphes 24 à 26 et 73 ; Lignes directrices du CEPD n° 2/2019 sur le traitement des données à caractère personnel au titre de l'article 6(1)(b) du RGPD dans le cadre de la fourniture de services en ligne aux personnes concernées, version 2.0, adoptées le 8 octobre 2019, paragraphes 23 à 25 ; Avis 11/2024 du CEPD sur l'utilisation de la reconnaissance faciale pour rationaliser le flux de passagers dans les aéroports, version 1.1, adopté le 23 mai 2024, paragraphe 27.

[57] Voir CJUE, arrêt du 16 décembre 2008, affaire C 524/06, Heinz Huber contre République fédérale d'Allemagne (ECLI:EU:C:2008:724), point 66. Voir également dans la même affaire les conclusions de l'avocat général Poiares Maduro dans l'affaire C-524/06, Heinz Huber contre Bundesrepublik Deutschland (ECLI:EU:C:2008:194), point 16, selon lesquelles: «le critère approprié ici est celui de l'efficacité, et il appartient à la juridiction nationale de l'appliquer. La question qu'elle doit se poser est de savoir s'il existe d'autres moyens de traitement des données par lesquels les autorités d'immigration pourraient faire respecter les règles relatives au statut de séjour. Si elle répond à cette question par l'affirmative, le stockage et le traitement centralisés des données des citoyens de l'Union devraient être déclarés illégaux. Il n'est pas nécessaire que le système alternatif soit le plus efficace ou le plus approprié; il suffit qu'il soit en mesure de fonctionner de manière adéquate. En d'autres termes, même si le registre central est plus efficace, plus pratique ou plus convivial que ses alternatives (comme les registres décentralisés et locaux), ces derniers sont clairement à préférer s'ils peuvent être utilisés pour indiquer le statut de résidence des citoyens de l'Union.

[58] Voir CJUE, arrêt du 27 septembre 2017, affaire C-73/16, Peter Puškár (ECLI:EU:C:2017:725), point 113 :
« Il appartient donc à la juridiction nationale de vérifier si l'établissement de la liste litigieuse et l'inscription des noms des personnes concernées dans un tel registre sont propres à atteindre les objectifs poursuivis par celles-ci et

73. Par exemple, et selon le cas, le volume de données personnelles prévu dans le modèle d'IA doit être évalué à la lumière d'alternatives moins intrusives qui peuvent raisonnablement être disponibles pour atteindre tout aussi efficacement l'objectif de l'intérêt légitime poursuivi. Si la poursuite de l'objectif est également possible grâce à un modèle d'IA qui n'implique pas le traitement de données personnelles, le traitement de données personnelles doit alors être considéré comme non nécessaire. Cela est particulièrement pertinent pour le développement
des modèles d'IA. Lorsqu'elles évaluent si la condition de nécessité est remplie, les autorités de surveillance doivent accorder une attention particulière à la quantité de données à caractère personnel traitées et à la question de savoir si elle est proportionnée à la poursuite de l'intérêt légitime en jeu, également à la lumière du principe de minimisation des données.

74. L'évaluation de la nécessité doit également tenir compte du contexte plus large du traitement envisagé des données à caractère personnel. L'existence de moyens moins intrusifs pour les droits et libertés fondamentaux des personnes concernées peut varier selon que le responsable du traitement a une relation directe avec les personnes concernées (données de première partie) ou non (données de tiers). La CJUE a fourni certaines considérations à prendre en compte lors de l'analyse de la nécessité du traitement de données de première partie aux fins des intérêts légitimes poursuivis (bien que dans le contexte de la divulgation de ces données à des tiers)59.

75. La mise en œuvre de mesures techniques de protection des données personnelles peut également contribuer à satisfaire au critère de nécessité. Il peut s'agir, par exemple, de mesures telles que celles identifiées à la section 3.2.2
de telle sorte que l'anonymisation ne soit pas atteinte, mais que cela réduise néanmoins la facilité avec laquelle les personnes concernées peuvent être identifiées. Le CEPD note que certaines de ces mesures, lorsqu'elles ne sont pas requises pour se conformer au RGPD, peuvent constituer des garanties supplémentaires, comme analysé plus en détail dans la sous-section « mesures d'atténuation », dans la section 3.3.2.360.

3.3.2.3 Troisième étape – Test d'équilibrage

76. La troisième étape de l'évaluation de l'intérêt légitime est l '« exercice de mise en balance » (également appelé dans le présent document « test de mise en balance »)61. Cette étape consiste à identifier et décrire les différents droits et intérêts opposés en jeu62, c'est-à-dire d'un côté les intérêts, les droits fondamentaux et les libertés des personnes concernées, et de l'autre côté les intérêts du responsable du traitement ou d'un tiers. Les circonstances spécifiques de l'affaire doivent ensuite être prises en considération pour démontrer que l'intérêt légitime constitue une base juridique appropriée pour les activités de traitement en jeu63.

« s'il n'existe pas d'autres moyens moins restrictifs pour atteindre ces objectifs » ; Voir également par exemple l'avis de l'avocat général Rantos dans l'affaire C-252/21, Meta contre Bundeskartellamt, ECLI:EU:C:2022:704, point 61, déclarant : « [...] Il est donc nécessaire qu'un lien étroit existe entre le traitement et l'intérêt poursuivi, en l'absence d'alternatives plus respectueuses de la protection des données, car il ne suffit pas que le traitement soit simplement utile au responsable du traitement ».

59 CJUE, arrêt du 4 octobre 2024, affaire C 621/22, Royal Dutch Lawn Tennis Association (ECLI:EU:C:2024:857), paragraphs 51-53.

60 Voir les lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 57.

61 Voir les lignes directrices du CEPD 1/2024 sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphes 31 à 60.

62 Voir les lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 32.

63 Voir les lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 32, faisant également référence à la CJUE, arrêt du 4 juillet 2023, affaire C-252/21, Meta c. Office fédéral des cartels (ECLI:EU:C:2023:537), point 110.

Intérêts, droits et libertés fondamentaux des personnes concernées

77. L'article 6(1)(f) du RGPD prévoit que, lors de l'évaluation des différentes composantes dans le cadre du test de mise en balance, le responsable du traitement doit tenir compte des intérêts, des droits fondamentaux et des libertés des personnes concernées. Les intérêts des personnes concernées sont ceux qui peuvent être affectés par le traitement en question. Français Dans le cadre de la phase de développement d'un modèle d'IA, ceux-ci peuvent inclure, sans s'y limiter, l'intérêt à l'autodétermination et à la conservation du contrôle de ses propres données personnelles (par exemple les données collectées pour le développement du modèle). Dans le cadre du déploiement d'un modèle d'IA, les intérêts des personnes concernées peuvent inclure, sans s'y limiter, l'intérêt à conserver le contrôle de ses propres données personnelles (par exemple les données traitées une fois le modèle déployé), des intérêts financiers (par exemple lorsqu'un modèle d'IA est utilisé par la personne concernée pour générer des revenus, ou est utilisé par un individu dans le cadre de son activité professionnelle), des avantages personnels (par exemple lorsqu'un modèle d'IA est utilisé pour améliorer l'accessibilité à certains services), ou des intérêts socio-économiques (par exemple lorsqu'un modèle d'IA permet l'accès à de meilleurs soins de santé, ou facilite l'exercice d'un droit fondamental tel que l'accès à l'éducation)64.

78. Plus un intérêt est défini avec précision à la lumière de la finalité poursuivie par le traitement, mieux c'est. elle permettra d'appréhender clairement la réalité des bénéfices et des risques à prendre en compte dans le test de mise en balance.

79. En ce qui concerne les droits et libertés fondamentaux des personnes concernées, le développement et le déploiement de modèles d'IA peuvent entraîner de graves risques pour les droits protégés par la Charte des droits fondamentaux de l'Union européenne (la « Charte de l'UE »), notamment le droit à la vie privée et familiale (article 7 de la Charte de l'UE) et le droit à la protection des données à caractère personnel (article 8 de la Charte de l'UE). Ces risques peuvent survenir pendant la phase de développement, par exemple lorsque des données à caractère personnel sont collectées pour contrer les droits des personnes concernées. Ces risques peuvent également survenir lors de la phase de déploiement, par exemple lorsque des données personnelles sont traitées par (ou dans le cadre de) le modèle d'une manière qui contrevient aux droits des personnes concernées, ou lorsqu'il est possible de déduire, accidentellement ou par des attaques (par exemple, inférence d'appartenance, extraction ou inversion de modèle), quelles données personnelles sont contenues dans la base de données d'apprentissage. De telles situations présentent un risque pour la vie privée des personnes concernées dont les données pourraient apparaître lors de la phase de déploiement du système d'IA (par exemple, risque de réputation, vol ou fraude d'identité, risque de sécurité en fonction de la nature des données).

80. Selon le cas, d'autres droits fondamentaux peuvent également être menacés. Par exemple, la collecte de données à grande échelle et sans discrimination par des modèles d'IA au cours de la phase de développement peut créer un sentiment de surveillance chez les personnes concernées, en particulier compte tenu des difficultés à empêcher le piratage des données publiques. Cela peut conduire les individus à s'autocensurer et présenter des risques de porter atteinte à leur liberté d'expression (article 11 de la Charte de l'UE). Dans la phase de déploiement, des risques pour la liberté d'expression sont également présents lorsque des modèles d'IA sont utilisés pour bloquer la publication de contenus par les personnes concernées. En outre, un modèle d'IA recommandant des contenus inappropriés à des personnes vulnérables peut présenter des risques pour leur santé mentale (article 3(1) de la Charte de l'UE). Dans d'autres cas, le déploiement de modèles d'IA peut également entraîner des conséquences néfastes sur le droit de l'individu à travailler (article 15 de la Charte de l'UE), par exemple lorsque les candidatures à un emploi sont présélectionnées à l'aide d'un modèle d'IA. De la même manière, un modèle d'IA pourrait présenter des risques pour le droit à la non-discrimination (article 21 de la Charte de l'UE), s'il est discriminatoire. individus en fonction de certaines caractéristiques personnelles (telles que la nationalité ou le sexe). En outre,

64 Voir les lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 38.

Le déploiement de modèles d'IA peut également présenter des risques pour la sécurité et la sûreté de l'individu (par exemple lorsque le modèle d'IA est utilisé avec une intention malveillante), ainsi que des risques pour son intégrité physique et mentale65.

81. Le déploiement de modèles d'IA peut également avoir un impact positif sur certains droits fondamentaux, par exemple le modèle peut soutenir le droit à l'intégrité mentale de la personne (article 3 de la Charte), par exemple lorsqu'un modèle d'IA est utilisé pour identifier des contenus préjudiciables en ligne ; ou le modèle peut faciliter l'accès à certains services essentiels ou faciliter l'exercice de droits fondamentaux, tels que l'accès à l'information (article 11 de la Charte de l'UE) ou l'accès à l'éducation (article 14 de la Charte de l'UE).

Impact du traitement sur les personnes concernées

82. Le traitement des données personnelles qui a lieu lors du développement et du déploiement de modèles d'IA peut avoir des répercussions différentes sur les personnes concernées, qui peuvent être positives ou négatives66. Par exemple, si une
Si l'activité de traitement comporte des avantages pour la personne concernée, ceux-ci peuvent être pris en compte dans le test de mise en balance. Si l'existence de tels avantages peut conduire une AS à conclure que les intérêts du responsable du traitement ou d'un tiers ne sont pas supplantés par les intérêts, les droits fondamentaux et les libertés des personnes concernées, une telle conclusion ne peut être que le résultat d'une analyse au cas par cas prenant en considération tous les facteurs appropriés.

83. L'impact du traitement sur les personnes concernées peut être influencé par (i) la nature des données traitées par les modèles ; (ii) le contexte du traitement ; et (iii) les conséquences supplémentaires que le traitement peut avoir67.

84. En ce qui concerne la nature des données traitées, il convient de rappeler que - outre les catégories particulières de données à caractère personnel et les données relatives aux condamnations pénales et aux infractions qui bénéficient respectivement d'une protection supplémentaire en vertu des articles 9 et 10 du RGPD - le traitement de certaines autres catégories de données
Les données personnelles peuvent avoir des conséquences importantes pour les personnes concernées. Dans ce contexte, le traitement de certains types de données personnelles révélant des informations hautement privées (par exemple des données financières ou des données de localisation) pour le développement et le déploiement d'un modèle d'IA doit être considéré comme pouvant avoir un impact grave sur les personnes concernées. Dans la phase de déploiement, les conséquences d'un tel traitement pour les personnes concernées peuvent par exemple être économiques (par exemple, discrimination dans le contexte de l'emploi) et/ou réputationnelles (par exemple, diffamation).

85. En ce qui concerne le contexte du traitement, il est d'abord nécessaire d'identifier les éléments qui pourraient créer des risques pour les personnes concernées (par exemple, la manière dont le modèle a été développé, la manière dont le modèle peut être déployé et/ou si les mesures de sécurité utilisées pour protéger les données personnelles
(sont appropriées). La nature du modèle et les utilisations opérationnelles prévues jouent un rôle clé dans l'identification de ces causes potentielles.

86. Il est également nécessaire d'évaluer la gravité de ces risques pour les personnes concernées. On peut notamment prendre en compte la manière dont les données personnelles sont traitées (par exemple si elles sont combinées avec d'autres ensembles de données), quelle est l'ampleur du traitement et la quantité de données personnelles traitées68 (par exemple, le volume global de données, le volume de données par personne concernée, le nombre de personnes concernées)69,

---

65 Lignes directrices 1/2024 sur le traitement des données personnelles basé sur l'article 6(1)(f) GDPR, version 1.0, adoptée le 8 octobre 2024, paragraphe 46.

66 Voir les lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 39.

67 Voir les lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 32.

68 Voir les lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 43.

69 CJUE, arrêt du 4 juillet 2023, affaire C-252/21, Meta contre Bundeskartellamt (ECLI:EU:C:2023:537), paragraphe 116.

le statut de la personne concernée (par exemple, enfants ou autres personnes vulnérables) et sa relation avec le responsable du traitement (par exemple, si la personne concernée est un client). Par exemple, l'utilisation du web scraping dans la phase de développement peut entraîner - en l'absence de garanties suffisantes - des impacts significatifs sur les personnes, en raison du volume important de données collectées, du grand nombre de personnes concernées et de la collecte indiscriminée de données à caractère personnel.

87. Les conséquences supplémentaires que le traitement peut entraîner doivent également être prises en compte lors de l'évaluation de l'impact du traitement sur les personnes concernées. Elles doivent être évaluées par les autorités de surveillance au cas par cas, compte tenu des faits précis en question.

88. Ces conséquences peuvent inclure (sans toutefois s'y limiter) des risques de violation des droits fondamentaux des personnes concernées, tels que décrits dans la sous-section précédente70. Les risques peuvent varier en probabilité et en gravité, et peuvent résulter d'un traitement de données à caractère personnel susceptible d'entraîner des dommages physiques, matériels ou immatériels, en particulier lorsque le traitement peut donner lieu à une discrimination71.

89. Lorsque le déploiement d'un modèle d'IA implique le traitement de données à caractère personnel à la fois (i) des personnes concernées dont les données à caractère personnel sont incluses dans l'ensemble de données utilisé dans la phase de développement ; et (ii) des personnes concernées dont les données à caractère personnel sont traitées dans la phase de déploiement, les autorités de surveillance devraient distinguer et prendre en compte les risques affectant les intérêts, les droits et les libertés de chacune de ces catégories de personnes concernées lors de la vérification du test de mise en balance effectué par un responsable du traitement.

90. Enfin, l'analyse des éventuelles conséquences supplémentaires du traitement doit également tenir compte de la probabilité que ces conséquences supplémentaires se matérialisent. L'évaluation de cette probabilité doit être effectuée en tenant compte des mesures techniques et organisationnelles en place et des circonstances spécifiques de l'affaire. Par exemple, les autorités de contrôle peuvent examiner si des mesures ont été mises en œuvre pour éviter une éventuelle utilisation abusive du modèle d'IA. Pour les modèles d'IA qui peuvent être déployés à diverses fins, comme l'IA générative, cela peut inclure des contrôles limitant autant que possible leur utilisation à des fins nuisibles, par exemple : la création de deepfakes ; les chatbots utilisés à des fins de désinformation,

phishing et autres types de fraudes ; et IA manipulatrice/agents IA (en particulier lorsqu'ils sont anthropomorphes ou fournissent des informations trompeuses).

Attentes raisonnables des personnes concernées

91. Sur la base du considérant 47 du RGPD, « [e]n tout état de cause, l'existence d'un intérêt légitime nécessiterait une évaluation minutieuse, notamment pour déterminer si une personne concernée peut raisonnablement s'attendre, au moment et dans le contexte de la collecte des données à caractère personnel, à ce qu'un traitement à cette fin puisse avoir lieu. Les intérêts et les droits fondamentaux de la personne concernée pourraient notamment prévaloir sur l'intérêt du responsable du traitement lorsque les données à caractère personnel sont traitées dans des circonstances où les personnes concernées ne s'attendent pas raisonnablement à un traitement ultérieur » 72.

92. Les attentes raisonnables jouent un rôle clé dans le test d'équilibrage, notamment en raison de la complexité de la technologie utilisée dans les modèles d'IA et du fait qu'il peut être difficile pour les personnes concernées de comprendre les attentes raisonnables.

---

70 Voir la sous-section « Intérêts, droits fondamentaux et libertés des personnes concernées » ci-dessus.

71 Voir la section 2.3 des lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, Version 1.0, adoptée le 8 octobre 2024. Voir également le considérant 75 du RGPD pour d'autres exemples.

72 Voir également CJUE, arrêt du 4 juillet 2023, affaire C-252/21, Meta c. Bundeskartellamt (ECLI:EU:C:2023:537), point 112 ; CJUE, arrêt du 11 décembre 2019, affaire C-708/18, Asociatia de Proprietari bloc M5A-ScaraA (ECLI:EU:C:2019:1064), point 58 ; CJUE, arrêt du 4 octobre 2024, affaire C-621/22, Royal Dutch Lawn Tennis Association (ECLI:EU:C:2024:857), point 55.

diversité des utilisations potentielles d'un modèle d'IA et du traitement des données impliqué73. À cette fin, les informations fournies aux personnes concernées peuvent être prises en compte pour évaluer si les personnes concernées peuvent raisonnablement s'attendre à ce que leurs données personnelles soient traitées. Cependant, si l'omission d'informations peut contribuer à ce que les personnes concernées ne s'attendent pas à un certain traitement, le simple respect des exigences de transparence énoncées dans le RGPD ne suffit pas en soi à considérer que les personnes concernées peuvent raisonnablement s'attendre à un certain traitement74. En outre, le simple fait que des informations relatives à la phase de développement d'un modèle d'IA soient incluses dans la politique de confidentialité du responsable du traitement ne signifie pas nécessairement que les personnes concernées peuvent raisonnablement s'attendre à un certain traitement on peut raisonnablement s'attendre à ce que cela se produise ; cela devrait plutôt être analysé par les AS en fonction des circonstances spécifiques de l'affaire et en tenant compte de tous les facteurs pertinents.

93. Lors de l'évaluation des attentes raisonnables des personnes concernées par rapport au traitement qui a lieu en phase de développement, il est important de se référer aux éléments mentionnés dans les lignes directrices du CEPD sur l'intérêt légitime75. En outre, dans le cadre de l'objet du présent avis, il est important de prendre en compte le contexte plus large du traitement. Cela peut inclure, sans toutefois s'y limiter, le fait que les données à caractère personnel étaient ou non accessibles au public, la nature de la relation entre la personne concernée et le responsable du traitement (et s'il existe un lien entre les deux), la nature du service, le contexte dans lequel les données à caractère personnel ont été collectées, la source à partir de laquelle les données ont été collectées (par exemple, le site Web ou le service où les données à caractère personnel ont été collectées et les paramètres de confidentialité qu'ils offrent), les utilisations potentielles ultérieures du modèle et le fait que les personnes concernées soient réellement conscientes que leurs données à caractère personnel sont en ligne.

94. Dans la phase de développement du modèle, les attentes raisonnables des personnes concernées peuvent différer selon que les données traitées pour développer le modèle sont rendues publiques par les personnes concernées.
ou non. En outre, les attentes raisonnables peuvent également différer selon que les données ont été directement fournies au responsable du traitement (par exemple dans le cadre de leur utilisation du service) ou que le responsable du traitement les a obtenues d'une autre source (par exemple via un tiers ou par scraping). Dans les deux cas, les mesures prises pour informer les personnes concernées des activités de traitement doivent être prises en compte lors de l'évaluation des attentes raisonnables.

95. Dans la phase de déploiement du modèle d'IA, il est tout aussi important de prendre en compte les attentes raisonnables des personnes concernées dans le contexte des capacités spécifiques du modèle. Par exemple, pour les modèles d'IA qui peuvent s'adapter en fonction des données fournies, il peut être pertinent de déterminer si les personnes concernées étaient conscientes qu'elles avaient fourni des données à caractère personnel afin que le modèle d'IA puisse ajuster ses paramètres.
des réponses à leurs besoins et afin qu'ils puissent obtenir des services sur mesure. En outre, il peut également être pertinent de déterminer si cette activité de traitement n'aurait d'impact que sur le service fourni aux personnes concernées (par exemple, la personnalisation du contenu pour un utilisateur spécifique) ou si elle serait utilisée pour modifier le service fourni à tous les clients (par exemple, pour améliorer le modèle de manière générale). Comme au stade du développement, il peut également être particulièrement pertinent de déterminer s'il existe un lien direct entre les personnes concernées et le responsable du traitement. Un tel lien direct peut, par exemple, permettre au responsable du traitement

---

[73] Par exemple, dans l'arrêt du 4 juillet 2023, affaire C-252/21, Meta contre Bundeskartellamt (ECLI:EU:C:2023:537), paragraphe 123, si la CJUE a estimé que « l'amélioration du produit » ne peut en principe être exclue en tant qu'intérêt légitime, elle a également estimé qu'il est « douteux que [...] l'objectif d'« amélioration du produit », compte tenu de l'ampleur de cet intérêt légitime, puisse être atteint ».
« le traitement et son impact significatif sur l'utilisateur, ainsi que le fait que l'utilisateur ne peut raisonnablement s'attendre à ce que ces données soient traitées [...] peuvent prévaloir sur les intérêts et les droits fondamentaux d'un tel utilisateur, notamment dans le cas où cet utilisateur est un enfant ».

[74] Lignes directrices 1/2024 sur le traitement des données personnelles basé sur l'article 6(1)(f) GDPR, version 1.0, adoptées le 8 octobre 2024, paragraphe 53.

[75] Lignes directrices 1/2024 sur le traitement des données personnelles fondées sur l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 Octobre 2024, paragraphes 50-54.

de fournir facilement des informations aux personnes concernées sur l'activité de traitement et le modèle, ce qui pourrait ensuite influencer les attentes raisonnables de ces personnes concernées.

Mesures d'atténuation

96. Lorsque les intérêts, droits et libertés des personnes concernées semblent prévaloir sur les intérêts légitimes poursuivis par le responsable du traitement ou un tiers, le responsable du traitement peut envisager d'introduire des mesures d'atténuation pour limiter l'impact du traitement sur ces personnes concernées. Les mesures d'atténuation sont des garanties qui doivent être adaptées aux circonstances de l'espèce et dépendent de différents facteurs, notamment de l'utilisation prévue du modèle d'IA. Ces mesures d'atténuation viseraient à garantir que les intérêts du responsable du traitement ou du tiers ne seront pas outrepassés, de sorte que le responsable du traitement puisse s'appuyer sur cette base juridique.

97. Comme le rappellent les lignes directrices du CEPD sur l'intérêt légitime, les mesures d'atténuation ne doivent pas être confondues avec les mesures que le responsable du traitement est légalement tenu d'adopter de toute façon pour garantir le respect du RGPD, que le traitement soit ou non fondé sur l'article 6(1)(f) du RGPD[76]. Cela est particulièrement important pour les mesures qui, par exemple, nécessitent de se conformer aux principes du RGPD, tels que le principe de minimisation des données.

98. La liste des mesures ci-dessous n'est ni exhaustive ni prescriptive et la mise en œuvre des mesures doit être envisagée au cas par cas. Bien que, selon les circonstances, certaines des mesures ci-dessous puissent être requises pour se conformer à des obligations spécifiques du RGPD, lorsque ce n'est pas le cas, elles peuvent être prises en compte en tant que garanties supplémentaires. En outre, certaines des mesures mentionnées ci-dessous concernent des domaines qui sont sujets à une évolution rapide et à de nouveaux développements, et doivent être prises en compte par les autorités de surveillance lorsqu'elles traitent un cas spécifique.

99. En ce qui concerne la phase de développement des modèles d'IA, plusieurs mesures peuvent être prises pour atténuer les risques posés par le traitement des données de première partie et de tiers (y compris pour atténuer les risques liés aux pratiques de scraping Web). Sur la base de ce qui précède, le CEPD fournit quelques exemples de mesures qui peuvent être mises en œuvre pour atténuer les risques identifiés dans le test de mise en balance, et qui devraient être pris en compte par les AS lors de l'évaluation de modèles d'IA spécifiques au cas par cas.

100. Mesures techniques :

a. Les mesures mentionnées à la section 3.2.2 qui sont appropriées pour atténuer les risques en jeu, lorsque ces mesures n'entraînent pas l'anonymisation du modèle et ne sont pas requises pour se conformer à d'autres obligations du RGPD ou au test de nécessité (deuxième étape de l'évaluation de l'intérêt légitime).

101. Outre celles-ci, d'autres mesures pertinentes peuvent inclure :

b. Mesures de pseudonymisation : il peut s'agir, par exemple, de mesures visant à empêcher toute combinaison de données fondée sur des identifiants individuels. Ces mesures peuvent ne pas être appropriées lorsque l'autorité de contrôle considère que le responsable du traitement a démontré la nécessité raisonnable de recueillir des données différentes sur une personne particulière pour le développement du système ou du modèle d'IA en question.

c. Mesures visant à masquer les données personnelles ou à les remplacer par de fausses données personnelles dans l'ensemble de formation (par exemple, le remplacement des noms et des adresses e-mail par de faux noms et de fausses adresses e-mail).

[76] Lignes directrices 1/2024 sur le traitement des données personnelles basé sur l'article 6(1)(f) GDPR, version 1.0, adoptées le 8 octobre 2024, paragraphe 57.

(adresses). Cette mesure peut être particulièrement appropriée lorsque le contenu substantiel réel des données n'est pas pertinent pour le traitement global (par exemple dans le cadre d'une formation LLM).

102. Mesures facilitant l'exercice des droits des personnes :

a. Observer une période de temps raisonnable entre la collecte d'un ensemble de données de formation et son utilisation. Cette garantie supplémentaire peut permettre aux personnes concernées d'exercer leurs droits pendant cette période, le délai raisonnable étant évalué en fonction des circonstances de chaque cas.
cas.

b. Proposer dès le départ une « option de retrait » inconditionnelle, par exemple en accordant aux personnes concernées un droit discrétionnaire de s'opposer avant que le traitement n'ait lieu, afin de renforcer le contrôle des individus sur leurs données, ce qui va au-delà des conditions de l'article 21 du RGPD77.

c. Permettre aux personnes concernées d'exercer leur droit à l'effacement même lorsque les motifs spécifiques énumérés à l'article 17(1) du RGPD ne s'appliquent pas78.

d. Permettre aux personnes concernées de soumettre des réclamations concernant la régurgitation ou la mémorisation de données à caractère personnel et les circonstances et moyens par lesquels les réclamations peuvent être reproduites, permettre aux responsables du traitement de reproduire et d'évaluer les techniques de désapprentissage pertinentes pour répondre aux réclamations.

103. Mesures de transparence : dans certains cas, les mesures d'atténuation pourraient inclure des mesures qui assurent une plus grande transparence en ce qui concerne le développement du modèle d'IA. Certaines mesures, en plus du respect des obligations du RGPD, peuvent contribuer à surmonter l'asymétrie d'information et permettre aux personnes concernées de mieux comprendre le traitement impliqué dans la phase de développement :

a. Diffusion de communications publiques et facilement accessibles qui vont au-delà des informations requises en vertu des articles 13 ou 14 du RGPD, par exemple en fournissant des détails supplémentaires sur les critères de collecte et tous les ensembles de données utilisés, en tenant compte de la protection particulière des enfants et des personnes vulnérables.

b. Formes alternatives d'information des personnes concernées, par exemple : campagnes médiatiques auprès de différents médias pour informer les personnes concernées, campagne d'information par courrier électronique, utilisation de visualisations graphiques, questions fréquemment posées, étiquettes de transparence et fiches modèles dont la systématisation pourrait structurer la présentation des informations sur les modèles d'IA, et
rapports annuels de transparence sur une base volontaire.

104. Mesures d'atténuation spécifiques dans le contexte du web scraping : Considérant que, comme mentionné ci-dessus, le web scraping présente des risques spécifiques79, des mesures d'atténuation spécifiques pourraient être identifiées dans ce contexte.
Le cas échéant, elles peuvent être prises en compte par les autorités de surveillance, en plus des mesures d'atténuation mentionnées ci-dessus, lors de l'enquête sur les responsables du traitement effectuant du scraping Web.

105. Des mesures spécifiques, lorsqu'elles ne sont pas nécessaires dans le cadre de la deuxième étape de l'évaluation de l'intérêt légitime, peuvent s'avérer utiles pour atténuer le risque dans le contexte du scraping Web. Il peut s'agir de mesures techniques, telles que :

77 Ibid.

78 Ibid.

79 Ces pratiques peuvent également soulever des questions supplémentaires qui ne sont pas abordées dans le présent avis. Voir par exemple Pagallo U., Ciani Sciolla J., Anatomy of web data scraping: ethical, standards, and the troubles of the law. Revue européenne du droit et des technologies de la vie privée, (2023) 2 p. 1 - 19, disponible à l'adresse suivante : https://doi.org/10.57230/EJPLT232PS.

a. Exclure le contenu des données des publications qui pourraient inclure des données personnelles comportant des risques pour des personnes particulières ou des groupes de personnes (par exemple, des personnes qui pourraient être soumises à des abus, à des préjugés ou même à des dommages corporels si les informations étaient rendues publiques).

b. Veiller à ce que certaines catégories de données ne soient pas collectées ou à ce que certaines sources soient exclues de la collecte de données; il peut s'agir, par exemple, de certains sites Web qui sont particulièrement intrusifs en raison de la sensibilité de leur sujet.

c. Exclure la collecte à partir de sites Web (ou de sections de sites Web) qui s'opposent clairement au web scraping et à la réutilisation de leur contenu à des fins de création de bases de données de formation d'IA (par exemple, en respectant les fichiers robots.txt ou ai.txt ou tout autre mécanisme reconnu pour exprimer l'exclusion de l'exploration ou du scraping automatisé).

d. Imposer d'autres limites pertinentes à la collecte, y compris éventuellement des critères basés sur des périodes de temps.

106. Dans le cadre du web scraping, des exemples de mesures spécifiques facilitant l'exercice des droits des personnes et la transparence peuvent inclure : la création d'une liste d'exclusion, gérée par le responsable du traitement et qui permet aux personnes concernées de s'opposer à la collecte de leurs données sur certains sites Web ou plateformes en ligne en fournissant des informations qui les identifient sur ces sites Web, y compris avant la collecte des données80.

107. Considérations spécifiques concernant les mesures d'atténuation dans la phase de déploiement : Bien que certaines des mesures mentionnées ci-dessus puissent également être pertinentes pour la phase de déploiement, selon les circonstances, le CEPD fournit ci-dessous une liste non exhaustive de mesures de soutien supplémentaires qui peuvent être mises en œuvre et qui devraient être évaluées par les autorités de surveillance au cas par cas.

a. Des mesures techniques peuvent par exemple être mises en place pour empêcher le stockage, la régurgitation ou la génération de données personnelles, notamment dans le cadre de modèles d'IA génératifs (tels que les filtres de sortie), et/ou pour atténuer le risque de réutilisation illégale par des modèles d'IA à usage général (par exemple, le tatouage numérique des sorties générées par l'IA).

b. Mesures facilitant ou accélérant l'exercice des droits des personnes dans la phase de déploiement, au-delà de ce qui est requis par la loi, concernant notamment, et sans s'y limiter, l'exercice du droit à l'effacement des données personnelles des données de sortie du modèle ou à la déduplication, et aux techniques post-formation qui tentent de supprimer ou de supprimer les données personnelles.

108. Lorsqu'elles examinent le déploiement d'un modèle d'IA spécifique, les autorités de contrôle doivent déterminer si le responsable du traitement a publié le test de mise en balance qu'il a effectué, car cela peut accroître la transparence et l'équité. Comme mentionné dans les lignes directrices du CEPD sur l'intérêt légitime, d'autres mesures peuvent être envisagées pour fournir aux personnes concernées des informations issues du test de mise en balance avant toute collecte de données à caractère personnel81. Le CEPD réitère également82 qu'un élément à prendre en compte est de savoir si le responsable du traitement a impliqué le DPD, le cas échéant.

80 Sauf si le responsable du traitement démontre qu'il existe des motifs légitimes et impérieux pour le traitement qui prévalent sur les intérêts et les droits et libertés de la personne concernée ou pour la constatation, l'exercice ou la défense de droits en justice.

81 Lignes directrices du CEPD 1/2024 sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 68.

82 Lignes directrices 1/2024 du CEPD sur le traitement des données à caractère personnel sur la base de l'article 6(1)(f) du RGPD, version 1.0, adoptées le 8 octobre 2024, paragraphe 12.

3.4 Sur l'impact éventuel d'un traitement illicite dans le cadre du développement d'un modèle d'IA sur la licéité du traitement ou de l'exploitation ultérieur du modèle d'IA

109. Cette section de l'avis répond à la question 4 de la demande. Cette question vise à obtenir des éclaircissements sur l'impact éventuel d'un traitement illicite dans la phase de développement sur le traitement ultérieur (par exemple dans la phase de déploiement du modèle d'IA) ou sur le fonctionnement du modèle. La question vise à aborder à la fois la situation dans laquelle un tel modèle d'IA traite des données à caractère personnel qui sont conservées dans le modèle (question 4(i) de la demande), ainsi que la situation dans laquelle aucun traitement de données à caractère personnel n'est plus impliqué dans le déploiement du modèle d'IA (c'est-à-dire que le modèle est anonyme) (question 4(ii) de la demande).

110. Avant d'aborder certains scénarios spécifiques, le CEPD fournit les considérations générales suivantes.

111. Premièrement, les éclaircissements apportés dans cette section porteront sur le traitement des données personnelles en phase de développement effectué en violation du principe de licéité tel qu'énoncé à l'article 5(1)(a) du RGPD et à l'article 6 du RGPD spécifiquement (ci-après « l'illicéité ») 83. Dans le même esprit, les considérations du CEPD porteront sur l'impact de l'illicéité du traitement en phase de développement sur la licéité (c'est-à-dire le respect de l'article 5(1)(a) du RGPD et de l'article 6 du RGPD) du traitement ou de l'exploitation ultérieurs du modèle. Toutefois, le CEPD souligne que le traitement effectué en phase de développement peut également conduire à des violations d'autres dispositions du RGPD, telles que le manque de transparence envers les personnes concernées, ou la protection des données dès la conception et/ou
défaut, qui ne sont pas analysés dans le présent avis.

112. Deuxièmement, lorsqu'on aborde cette question, le principe de responsabilité, qui exige que les responsables du traitement soient responsables et démontrent leur conformité avec, entre autres, l'article 5(1) du RGPD et l'article 6 du RGPD84, joue un rôle essentiel. Cela est également vrai pour la nécessité d'évaluer quelle organisation est responsable du traitement de l'activité de traitement en cause et si des situations de responsabilité conjointe se présentent (car elles peuvent être inextricablement liées)85. Compte tenu de l'importance des circonstances factuelles de chaque cas, y compris en ce qui concerne le rôle joué par chaque partie impliquée dans le traitement, les considérations du CEPD doivent être comprises comme des observations générales qui doivent être évaluées au cas par cas par SA.

113. Troisièmement, le CEPD souligne que, conformément à l'article 51(1) du RGPD, les autorités de contrôle sont « chargées de surveiller l'application du [RGPD], afin de protéger les droits et libertés fondamentaux des personnes physiques à l'égard du traitement et de faciliter la libre circulation des données à caractère personnel au sein de l'Union ». Il est donc de la compétence des autorités de contrôle d'évaluer la licéité du traitement et d'exercer les pouvoirs qui leur sont conférés par le RGPD conformément à leur cadre national86. Dans de tels cas, les autorités de contrôle disposent de pouvoirs discrétionnaires pour évaluer la ou les violations éventuelles et choisir les mesures appropriées et nécessaires

---

[83] CJUE, arrêt du 4 mai 2023, affaire C-60/22, Bundesrepublik Deutschland (ECLI:EU:C:2023:373), paragraphes 55-57.

[84] CJUE, arrêt du 4 mai 2023, affaire C-60/22, République fédérale d'Allemagne (ECLI:EU:C:2023:373), paragraphe 53.

[85] Lignes directrices du CEPD 07/2020 sur les concepts de responsable du traitement et de sous-traitant dans le RGPD, version 2.1, adoptées le 7 Juillet 2021, paragraphe 55.

[86] Des règles nationales spécifiques peuvent devoir être prises en compte. Voir par exemple l'article 2-decies du Code italien de protection des données (décret législatif 196/2003) qui établit que les données traitées en violation des règles de protection des données ne peuvent pas être utilisées. Ceci est sans préjudice d'autres cadres juridiques nationaux, tels que les lois pénales.

et des mesures proportionnées, parmi celles mentionnées à l'article 58 du RGPD, tenant compte des circonstances de chaque cas individuel87.

114. Lorsqu'une infraction est constatée, les autorités de contrôle peuvent imposer des mesures correctives, comme ordonner aux responsables du traitement, compte tenu des circonstances de chaque cas, de prendre des mesures pour remédier à l'illicéité du traitement initial. Il peut s'agir, par exemple, d'infliger une amende, d'imposer une limitation temporaire du traitement, d'effacer une partie de l'ensemble de données qui a été traité illégalement ou, lorsque cela n'est pas possible, en fonction des faits, compte tenu de la proportionnalité de la mesure, d'ordonner l'effacement de l'ensemble des données utilisées pour développer le modèle d'IA et/ou du modèle d'IA lui-même. Lors de l'évaluation de la proportionnalité de la mesure envisagée, les autorités de contrôle peuvent prendre en compte les mesures qui peuvent être appliquées par le responsable du traitement pour remédier à l'illicéité du traitement initial (par exemple, une nouvelle formation).

115. Le CEPD souligne également que, lorsque des données à caractère personnel sont traitées illégalement, les personnes concernées peuvent demander la suppression de leurs données à caractère personnel, sous réserve des conditions énoncées à l'article 17 du RGPD, et que les autorités de surveillance peuvent ordonner l'effacement des données à caractère personnel d'office88.

116. Lorsqu'elles évaluent si une mesure est appropriée, nécessaire et proportionnée, les autorités de surveillance peuvent prendre en compte, entre autres éléments, les risques encourus par les personnes concernées, la gravité de l'infraction, la faisabilité technique et financière de la mesure, ainsi que le volume de données personnelles concernées.

117. Enfin, le CEPD rappelle que les mesures prises par les autorités de surveillance en vertu du RGPD sont sans préjudice de celles prises par les autorités compétentes en vertu de la loi sur l'IA et/ou d'autres cadres juridiques applicables (par exemple législation sur la responsabilité civile).

118. Dans les sections suivantes, le CEPD abordera trois scénarios couverts par la question 4 de la Demande, lorsque les différences résident dans la question de savoir si les données personnelles traitées pour développer le modèle sont conservées dans le modèle, et/ou si le traitement ultérieur est effectué par le même responsable du traitement ou par un autre.

3.4.1 Scénario 1. Un responsable du traitement traite illégalement des données à caractère personnel pour développer le modèle, les données à caractère personnel sont conservées dans le modèle et sont ensuite traitées par le même responsable du traitement (par exemple dans le cadre du déploiement du modèle)

119. Ce scénario se rapporte à la question 4(i) de la demande, dans la situation où un responsable du traitement traite illégalement des données à caractère personnel (c'est-à-dire en ne respectant pas l'article 5(1)(a) du RGPD et l'article 6 du RGPD) pour développer un modèle d'IA, le modèle d'IA conserve des informations relatives à une personne physique identifiée ou identifiable et n'est donc pas anonyme. Les données à caractère personnel sont ensuite traitées ultérieurement par le même responsable du traitement (par exemple dans le cadre du déploiement du modèle). En ce qui concerne ce scénario, le CEPD fournit les considérations suivantes.

---

[87] Voir à cet égard considérant 129 du RGPD, ainsi que CJUE, arrêt du 26 septembre 2024, affaire C-768-21, TR contre Land Hessen (ECLI:EU:C:2024:785), paragraphe 37 ; Arrêt de la CJUE du 7 décembre 2023, dans les affaires jointes C-26/22 et C-64/22, SCHUFA Holding (Libération de reliquat de dette) (ECLI:EU:C:2023:958), point 57 ; et CJUE, arrêt du 14 mars 2024, affaire C-46/23, Újpesti Polgármester Hivatal (ECLI:EU:C:2024:239), point 34.

[88] À cet égard, l'avis 39/2021 du CEPD sur la question de savoir si l'article 58(2)(g) du RGPD pourrait servir de base juridique à une autorité de contrôle pour ordonner d'office l'effacement de données à caractère personnel dans une situation où une telle demande n'a pas été présentée par la personne concernée, paragraphe 28. Voir également, à cet égard, CJUE, arrêt du 14 mars 2024, affaire C-46/23, Újpesti Polgármesteri Hivatal (ECLI:EU:C:2024:239), paragraphe 42.

120. Le pouvoir de l'AS d'imposer des mesures correctives sur le traitement initial (comme expliqué aux paragraphes 113, 114, 115 ci-dessus) aurait en principe un impact sur le traitement ultérieur (par exemple si l'AS ordonne au responsable du traitement d'effacer les données à caractère personnel qui ont été traitées illégalement, ces mesures correctrices ne permettraient pas à ce dernier de traiter ultérieurement les données à caractère personnel qui ont fait l'objet des mesures).

121. En ce qui concerne spécifiquement l'impact du traitement illicite dans la phase de développement sur le traitement ultérieur (par exemple dans la phase de déploiement), le CEPD rappelle qu'il appartient aux autorités de sécurité de procéder à une analyse au cas par cas qui prend en compte les circonstances spécifiques de chaque cas.

122. La question de savoir si les phases de développement et de déploiement impliquent des finalités distinctes (constituant ainsi des activités de traitement distinctes) et la mesure dans laquelle l'absence de base juridique pour l'activité de traitement initiale affecte la licéité du traitement ultérieur doivent être évaluées au cas par cas, en fonction du contexte de l'affaire.

123. Par exemple, en ce qui concerne spécifiquement la base juridique de l'article 6(1)(f) du RGPD, lorsque le traitement ultérieur est fondé sur un intérêt légitime, le fait que le traitement initial était illicite devrait être pris en compte dans l'évaluation de l'intérêt légitime (par exemple en ce qui concerne les risques pour les personnes concernées ou le fait que les personnes concernées ne peuvent pas s'attendre à un tel traitement ultérieur). Dans ces cas, l'illicéité du traitement en phase de développement peut avoir une incidence sur la licéité du traitement ultérieur.

3.4.2 Scénario 2. Un responsable du traitement traite illégalement des données à caractère personnel pour développer le modèle, les données à caractère personnel sont conservées dans le modèle et sont traitées par un autre responsable du traitement dans le cadre du déploiement du modèle

124. Ce scénario se rapporte à la question 4(i) de la demande. Il diffère du scénario 1 (à la section 3.4.1 du présent avis) car les données à caractère personnel sont ensuite traitées par un autre responsable du traitement dans le cadre du déploiement du modèle d'IA.

125. Le CEPD rappelle que la détermination des rôles attribués à ces différents acteurs dans le cadre de la protection des données est une étape essentielle pour identifier les obligations qui s'appliquent en vertu du RGPD et qui est responsable de ces obligations, et que les situations de contrôle conjoint doivent également être prises en compte lors de l'évaluation des responsabilités de chaque partie en vertu du RGPD. Par conséquent, les observations ci-dessous doivent être considérées comme des éléments généraux qui devraient être pris en compte par les autorités de contrôle lorsque pertinent. En ce qui concerne ce scénario 2, le CEPD fournit les considérations suivantes.

126. Premièrement, il convient de rappeler que, conformément à l'article 5(1)(a) du RGPD, lu à la lumière de l'article 5(2) du RGPD, Chaque responsable du traitement doit s'assurer de la licéité du traitement qu'il effectue et être en mesure de le démontrer. Par conséquent, les autorités de contrôle doivent évaluer la licéité du traitement effectué par (i) le responsable du traitement qui a initialement développé le modèle d'IA ; et (ii) le responsable du traitement qui a acquis le modèle d'IA et traite lui-même les données à caractère personnel.

127. Deuxièmement, les considérations formulées aux paragraphes 113, 114 et 115 ci-dessus sont pertinentes dans la présente affaire, en ce qui concerne le pouvoir des autorités de contrôle d'intervenir en ce qui concerne le traitement initial. L'article 17(1)(d) du RGPD (effacement des données traitées de manière illicite) et l'article 19 du RGPD (obligation de notification concernant la rectification ou l'effacement des données à caractère personnel ou la limitation du traitement) peuvent, selon les circonstances de l'espèce, également être pertinents dans ce contexte, par exemple en ce qui concerne la notification que le responsable du traitement qui développe le modèle doit effectuer à l'égard du responsable du traitement qui déploie le modèle.

128. Troisièmement, en ce qui concerne l'impact éventuel de l'illicéité du traitement initial sur le traitement ultérieur effectué par un autre responsable du traitement, une telle évaluation devrait être effectuée par les autorités de contrôle au cas par cas.

129. Les autorités de surveillance devraient déterminer si le responsable du traitement qui déploie le modèle a procédé à une évaluation appropriée, dans le cadre de ses obligations de responsabilité89 pour démontrer la conformité avec l'article 5(1)(a) et l'article 6 du RGPD, afin de s'assurer que le modèle d'IA n'a pas été développé en traitant illégalement des données à caractère personnel. Cette évaluation par les autorités de surveillance devrait tenir compte du fait que le responsable du traitement a évalué certains critères non exhaustifs, tels que la source des données et le fait que le modèle d'IA résulte ou non d'une violation du RGPD, en particulier si cela a été déterminé par une autorité de surveillance ou un tribunal, de sorte que le responsable du traitement qui déploie le modèle ne puisse ignorer que le traitement initial était illicite.

130. Le responsable du traitement doit par exemple déterminer si les données proviennent d'une violation de données à caractère personnel ou si le traitement a fait l'objet d'une constatation d'infraction par une autorité de contrôle ou un tribunal. Le degré d'évaluation du responsable du traitement et le niveau de détail attendu par les autorités de contrôle peuvent varier en fonction de divers facteurs, notamment du type et du degré des risques soulevés par le traitement dans le modèle d'IA lors de son déploiement par rapport aux personnes concernées dont les données ont été utilisées pour développer le modèle.

131. Le CEPD note que la loi sur l'IA exige que les fournisseurs de systèmes d'IA à haut risque établissent une déclaration de conformité de l'UE90, et que cette déclaration contienne une déclaration selon laquelle le système d'IA concerné est conforme aux lois de l'UE sur la protection des données91. Le CEPD note qu'une telle auto-déclaration ne constitue pas nécessairement une constatation concluante de conformité au titre du RGPD. Elle peut néanmoins être prise en compte par les autorités de surveillance lors de l'examen d'un modèle d'IA spécifique.

132. Les mêmes considérations formulées au paragraphe 123 ci-dessus sont également pertinentes dans la présente affaire. Lorsque les autorités de contrôle vérifient si et comment le responsable du traitement a évalué le caractère approprié de l'intérêt légitime comme base juridique du traitement qu'il effectue, l'illicéité du traitement initial doit être prise en compte dans le cadre de l'évaluation de l'intérêt légitime, par exemple en évaluant les risques potentiels qui peuvent survenir pour les personnes concernées dont les données à caractère personnel ont été traitées illégalement pour développer le modèle. Différents aspects, soit de nature technique (par exemple l'existence de filtres ou de limitations d'accès mis en place lors du développement du modèle, que le responsable du traitement ultérieur ne peut ni contourner ni influencer, et qui pourraient empêcher l'accès aux données à caractère personnel ou leur divulgation) ou de nature juridique (par exemple la nature et la gravité de l'illicéité du traitement initial) doivent être dûment pris en considération dans le cadre du test de mise en balance.

3.4.3 Scénario 3. Un responsable du traitement traite illégalement des données à caractère personnel pour développer le modèle, puis s'assure que le modèle est anonymisé, avant que le même responsable du traitement ou un autre responsable du traitement n'initie un autre traitement de données à caractère personnel dans le cadre du déploiement

133. Ce scénario se rapporte à la question 4(ii) de la demande et fait référence à un cas dans lequel un responsable du traitement traite illégalement des données à caractère personnel pour développer le modèle d'IA, mais le fait d'une manière qui garantit que les données à caractère personnel sont anonymisées, avant que le même responsable du traitement ou un autre responsable du traitement n'initie un autre traitement de données à caractère personnel dans le cadre du déploiement. Tout d'abord, le CEPD rappelle que les autorités de contrôle sont compétentes et ont le pouvoir d'intervenir en ce qui concerne le traitement lié à l'anonymisation du modèle, ainsi que le traitement effectué pendant la phase de développement. Ainsi, les autorités de contrôle peuvent, en fonction des

[89] Article 5(2) du RGPD et article 24 du RGPD.
[90] Article 16(g) et article 47 de la loi AI.
[91] Annexe V, point 5 de la loi AI.

circonstances de l'espèce, imposer des mesures correctives à ce traitement initial (comme expliqué aux paragraphes 113, 114, 115 ci-dessus)

134. S'il peut être démontré que l'exploitation ultérieure du modèle d'IA n'entraîne pas le traitement
En ce qui concerne les données personnelles, le CEPD considère que le RGPD ne s'appliquerait pas92. Par conséquent, l'illicéité du traitement initial ne devrait pas avoir d'impact sur le fonctionnement ultérieur du modèle. Toutefois, le CEPD souligne qu'une simple affirmation de l'anonymat du modèle ne suffit pas à l'exempter de l'application du RGPD, et note que les autorités de contrôle devraient l'évaluer en tenant compte, au cas par cas, de la manière dont le modèle est traité. sur la base des considérations fournies par le CEPD pour répondre à la question 1 de la demande.

135. Lorsque les responsables du traitement traitent ultérieurement des données à caractère personnel collectées au cours de la phase de déploiement, après que le modèle a été anonymisé, le RGPD s'appliquerait à ces activités de traitement. Dans ces cas, en ce qui concerne le RGPD, la licéité du traitement effectué au cours de la phase de déploiement ne devrait pas être affectée par l'illicéité du traitement initial.

4 Remarques finales

136. Le présent avis s'adresse à toutes les AS et sera rendu public conformément à l'article 64(5)(b) du RGPD.

Pour le Comité européen de la protection des données

La chaise

Anu Talus

92 Considérant 26 RGPD.