

CONSULTATION PUBLIQUE

Fiches pratiques IA sur la mobilisation de l'intérêt légitime pour le développement de systèmes d'intelligence artificielle

Synthèse des contributions

Juin 2025

La CNIL a lancé, le 10 juin 2024, une consultation publique sur le développement de systèmes d'IA.

Les contributions ont nourri les travaux du deuxième lot de huit fiches pratiques, dont celles sur l'intérêt légitime pour leur publication définitive sur le site de la CNIL.

La synthèse en chiffres

Le 8 avril 2024, à la suite d'une consultation publique, **la CNIL a publié [un premier lot de fiches pratiques sur le développement des systèmes d'IA](#)** qui répondent en particulier aux interrogations portant sur l'application des principes de finalité, de minimisation et de durée de conservation pour le développement des systèmes d'IA et précisent certaines règles applicables à la recherche scientifique, à la réutilisation de bases de données ou à la réalisation d'analyse d'impact sur la protection des données (AIPD).

Dans la continuité de ces travaux, **la CNIL a lancé, le 10 juin 2024, [une nouvelle consultation publique sur un questionnaire concernant l'application du RGPD aux modèles d'IA et les 7 fiches suivantes](#)** :

- Base légale de l'intérêt légitime et développement de systèmes d'IA.
- Intérêt légitime : focus sur la diffusion des modèles en source ouverte (open source).
- Intérêt légitime : focus sur le moissonnage (*web moissonnage*).
- Informer les personnes concernées.
- Respecter et faciliter l'exercice des droits des personnes concernées.
- Annoter les données.
- Garantir la sécurité du développement d'un système d'IA.

À l'issue de la consultation publique, le 1^{er} octobre 2024, la CNIL a reçu 62 contributions représentatives de différents secteurs :

- des organismes à but lucratif :
 - 28 entreprises privées, dont des fournisseurs de systèmes d'IA, des utilisateurs de systèmes d'IA, ou encore d'autres entreprises de différents secteurs (IA, finance, santé, aéronautique, opérateurs de plateformes en ligne, publicité en ligne, audiovisuel, etc.) ;
 - 11 organisations représentantes des professionnels ;
 - 4 cabinets d'avocat ou de conseil.
- des organismes à but non-lucratif :
 - 3 associations représentatives de la société civile ;
 - 5 chercheurs ou groupement de chercheurs académiques et 1 think tank ;
- 6 particuliers ;
- 4 administrations publiques.

Cette synthèse se concentre sur les retours qui ont concerné les fiches « Mobiliser la base légale de l'intérêt légitime pour développer un système d'IA » et « La base légale de l'intérêt légitime : fiche focus sur les mesures à prendre en cas de collecte des données par moissonnage (*web moissonnage*) ».

Pour tenir compte des retours sur **la fiche n° 8 1/2 relative à la diffusion des modèles en source ouverte (*open source*)**, la CNIL prévoit de faire une nouvelle fiche « cas d'usage » qui fera l'objet d'une publication ultérieure. Cela permettra de préciser plus exhaustivement les conditions à respecter pour assurer la conformité de la diffusion en open source aux dispositions du RGPD.

Ces contributions ont permis à la CNIL :

- de faire évoluer ses projets de fiches en apportant des clarifications supplémentaires et consolider ses analyses ;
- d'apporter des réponses, dans la synthèse ci-dessous, aux préoccupations les plus fréquemment partagées par les contributeurs.

Fiche sur l'intérêt légitime

A. Remarques générales

1. Sur l'approche de la fiche

Synthèse des contributions

De nombreuses contributions ont demandé des clarifications ou des exemples supplémentaires pour rendre la fiche plus spécifique et opérationnelle.

Quelques contributions demandent des contenus plus spécifiques selon les sujets (une fiche spécifique à l'IA générative ou des recommandations dédiées au traitement de l'image et de la voix) ou selon les destinataires (par ex. des résumés exécutifs pour les décideurs).

Réponse de la CNIL

La CNIL a ajouté des exemples dans le corps de la fiche afin de préciser l'application des conditions de mobilisation de l'intérêt légitime à certains cas limite.

Elle réalisera également une synthèse des fiches sur la forme d'un contenu simplifié.

2. Questions générales sur l'intérêt légitime

Synthèse des contributions

Des questions générales ont été posées sans être spécifiques à l'IA (documentation de la validité de la base légale, mobilisation de l'intérêt légitime par une entité publique, etc.).

Réponse de la CNIL

La plupart de ces questions font l'objet de [contenus déjà existants publiés par la CNIL](#) ou par le CEPD, notamment dans le cadre du [projet de ligne directrices sur la base légale de l'intérêt légitime](#) en cours de finalisation suite à une consultation publique. Les interrogations des contributeurs seront remontées au CEPD pour alimenter ces travaux.

3. Questions non-spécifiques à l'intérêt légitime

Synthèse des contributions

Des questions ont été posées sur des points qui ne sont toutefois pas spécifiques à l'intérêt légitime. Par ailleurs, certains contributeurs s'interrogent sur l'harmonisation des recommandations à l'échelle européenne.

Réponse de la CNIL

La première série de fiches pratiques qui ont été publiées par la CNIL le 8 avril 2024 répondent à certaines interrogations, concernant notamment :

- **le périmètre de la fiche** : [voir la fiche introductive](#). Pour rappel, les fiches pratiques adoptées jusqu'à présent portent sur la phase de développement et non de déploiement. Les questions relatives au déploiement des modèles d'IA feront l'objet de publications ultérieures de la CNIL ;
- **l'applicabilité du RGPD au développement des modèles d'IA** : [voir la fiche introductive](#) ;
- **la qualification des acteurs et la distinction entre les rôles du développeur et du déployeur** : [voir la fiche pratique n°3](#). La CNIL rappelle que les obligations qui pèsent sur chacun des acteurs dépendent d'abord de leur qualification au regard du RGPD, qui doit être examinée au cas par cas ;
- **la validité des autres bases légales** : [voir la fiche pratique n°4](#) ;
- **la question de la finalité statistique des traitements** : [voir la fiche n° 4 2/2](#) ;
- **la distinction entre modèle et système** : [voir la synthèse des contributions](#).

Sur l'harmonisation des recommandations à l'échelle européenne, des travaux sont en cours au sein du CEPD. Celui-ci a récemment adopté [l'avis 28/2024 sur l'utilisation des données à caractère personnel pour le développement et le déploiement de modèles d'IA](#). De plus, la CNIL participe activement à l'élaboration des lignes directrices sur l'articulation entre le règlement sur l'intelligence artificielle (RIA) et le RGPD, des lignes directrices sur le moissonnage pour l'entraînement des modèles d'IA, ainsi que des codes de bonnes pratiques de la Commission européenne pour le RIA.

B. Sur l'articulation entre l'intérêt légitime et le consentement

Synthèse des contributions

De nombreuses contributions sont revenues sur la question de l'articulation entre le consentement et l'intérêt légitime avec deux positions divergentes :

- D'une part, **certains acteurs insistent sur l'absence de hiérarchie entre la base légale du consentement et de l'intérêt légitime** en se référant à la doctrine constante du CEPD. Ces contributions considèrent également que la possibilité de recueillir le consentement ne devrait pas avoir d'impact sur la possibilité de mobiliser l'intérêt légitime.
- À l'inverse, **des représentants de la société civile** critiquent le recours, par défaut, à l'intérêt légitime, **en considérant qu'il faudrait exiger le consentement** ou que le traitement pourrait également se fonder sur le contrat. Ces contributions soulignent en outre que les données des utilisateurs ayant donné leur consentement seraient suffisantes pour l'entraînement des modèles. Une contribution considère que **s'il est possible d'obtenir une autorisation (du titulaire de droits, de la personne concernée, etc.)**, il devrait être exigé d'obtenir le consentement.

Réponse de la CNIL

Il n'y a pas de hiérarchie entre les bases légales. Chaque base juridique est autonome et s'applique en fonction des circonstances, sans qu'aucune ne soit « supérieure » aux autres. Cette position est d'ailleurs confirmée par les [lignes directrices du CEPD sur l'intérêt légitime](#) en cours de finalisation suite à une consultation publique et [l'avis du CEPD sur les modèles d'IA](#).

La CNIL considère qu'en matière de développement des systèmes d'IA, la base légale de l'intérêt légitime peut être mobilisée sous réserve de mettre en place des garanties fortes qui permettent d'en respecter les conditions de validité.

En revanche, la CNIL précise dans la fiche que le consentement sera nécessaire lorsque les conditions de validité de l'intérêt légitime ne pourront pas être réunies. Elle indique qu'il pourra également être nécessaire au titre d'autres législations (par exemple, le DMA).

C. Sur la condition relative à la légitimité de l'intérêt

Synthèse des contributions

Plusieurs contributions saluent les exemples fournis par la CNIL comme étant des intérêts légitimes, notamment s'agissant de la recherche scientifique et des intérêts commerciaux. **Certaines contributions proposent à la CNIL d'être encore plus explicite sur la reconnaissance des intérêts commerciaux comme intérêts légitimes**, notamment au regard des divergences au sein de l'UE.

Plusieurs contributions **craignent que certains des critères** (notamment relatifs au caractère « réel et présent » de l'intérêt poursuivi ou de l'existence d'un lien entre l'intérêt et « la mission et l'activité de l'organisme ») **remettent en cause la possibilité de mener des travaux de R&D ou de développer de nouvelles activités**.

Enfin, plusieurs contributions (notamment des entreprises qui développent des modèles d'IA) s'interrogent sur la possibilité de considérer que **le « développement (et l'amélioration) d'un modèle d'IA à usage général » constitue un intérêt légitime**, indépendamment de l'usage qui en est fait. En effet :

- Certaines de ces contributions proposent **d'adopter la même approche que celle proposée par la CNIL pour définir la finalité du développement d'un système d'IA à usage général de**

manière suffisamment précise. De la même manière, elles souhaiteraient que l'intérêt puisse être défini au regard du type et des fonctionnalités du système d'IA développé.

- Certaines contributions de chercheurs et de représentants de la société civile soulignent à l'inverse que « *le développement d'un système d'IA* » ne peut pas être considéré comme un intérêt légitime in suffisamment précis et spécifique. Elles considèrent qu'il est nécessaire de **circonscrire davantage les intérêts qui peuvent être considérés comme légitimes.**

Réponse de la CNIL

La version finale de la fiche a été modifiée pour préciser qu'un intérêt commercial constitue un intérêt légitime pour autant qu'il ne soit pas contraire à la loi et que le traitement soit nécessaire et proportionné.

La CNIL rappelle que **ces critères ne font pas obstacle aux travaux de R&D, aux nouvelles activités, ni aux travaux qui ont un caractère expérimental.** En revanche, si les données sont réutilisées pour une finalité autre que celles pour quoi elles avaient été initialement collectées, il sera nécessaire d'obtenir le consentement de la personne ou d'effectuer un test de compatibilité (voir la [fiche n° 4 2/2 sur la réutilisation des données](#)).

Enfin, la CNIL a souligné dans sa fiche que le « *développement (et l'amélioration) d'un modèle d'IA à usage général* » doit être précisé avec, a minima, l'objectif poursuivi par le développement du modèle, en indiquant notamment si l'intérêt est d'ordre commercial/public/de recherche scientifique et pour des objectifs internes ou externes à la société.

D. Sur la nécessité du traitement

Synthèse des contributions

Si une contribution salue l'approche de cette partie, des contributions sont plutôt critiques au motif que la fiche ne fournit pas suffisamment de précision sur la manière de remplir la condition de nécessité.

Des contributions ont exprimé **des inquiétudes quant au risque que le critère de nécessité soit appliqué de manière trop restrictive ou trop flexible.** Une contribution considère notamment que l'utilisation de grands volumes de données pour le « développement d'IA » ne peut pas se fonder sur l'intérêt légitime.

Plusieurs contributions ont proposé ou s'interrogent sur **différents critères à prendre en compte pour évaluer la nécessité du traitement**, notamment l'utilisation du système et les évolutions technologiques.

Réponse de la CNIL

La CNIL a affirmé, dès le premier lot de fiches (voir notamment les fiches sur la minimisation), que **la réglementation ne s'opposait pas à l'utilisation de grands volumes de données qui peuvent être nécessaires pour l'entraînement d'IA.** Les données utilisées devront en revanche, en principe, avoir été sélectionnées pour optimiser l'entraînement de l'algorithme tout en évitant l'utilisation de données personnelles inutiles. Cette condition, en lien avec le principe de minimisation, implique aussi pour le responsable du traitement de privilégier la technique permettant d'atteindre le résultat recherché (ou du même ordre) en utilisant le moins de données personnelles possible. En particulier, le recours à l'apprentissage profond ne doit donc pas être systématique.

Sur les **critères à prendre en compte pour l'évaluation de la nécessité**, la CNIL rappelle que celle-ci doit s'apprécier au regard de l'intérêt poursuivi : lorsque l'utilisation du modèle d'IA est connue dès la phase de développement, ce sera un critère pertinent pour évaluer la nécessité du traitement.

La CNIL précise également, dans la fiche, que les évolutions technologiques doivent être prises en compte pour évaluer la nécessité de traiter un volume très important de données.

E. Sur la mise en balance

Remarques générales

- ❖ **Sur la clarification de l'analyse**

Synthèse des contributions

Certaines contributions **estiment impossible d'atteindre la mise en balance** pour le développement des systèmes d'IA, au regard de l'ampleur du traitement et de son impact en l'absence des attentes raisonnables.

Réponse de la CNIL

Il n'est pas possible de tirer de conclusion générale et catégorique sur la possibilité ou non de fonder le développement d'un système d'IA sur l'intérêt légitime. La CNIL rappelle, dans ses fiches, que l'analyse de la mise en balance doit s'effectuer au cas par cas, en tenant compte des circonstances spécifiques du traitement.

Elle a ajouté également des exemples permettant de préciser les cas dans lesquels la mise en balance pourrait ou non être considérée comme atteinte.

❖ Sur l'articulation avec d'autres outils d'analyse des risques

Synthèse des contributions

Plusieurs contributions s'interrogent sur la possibilité d'aligner l'analyse des risques dans le cadre de la mise en balance avec d'autres outils d'analyse des risques, dont l'AIPD et les analyses de risques prévues par le RIA.

Réponse de la CNIL

La validité de l'intérêt légitime doit pouvoir être démontrée pour tout développement d'un système d'IA qui implique le traitement de données personnelles. S'il peut être pertinent, dans certains cas, de prendre en compte les résultats d'autres analyses réalisées, notamment au titre des obligations du RIA, il ne faut toutefois pas confondre les différents exercices au risque de ne pas parvenir à démontrer la conformité du traitement aux conditions spécifiques de l'intérêt légitime.

Cette analyse peut toutefois s'articuler avec celles prévues par le RIA :

- lorsque le responsable du traitement est un fournisseur d'IA à haut risque au sens de l'article 6 du RIA ou un fournisseur d'un modèle d'IA à usage général présentant des risques systémiques au sens de l'article 51 du RIA, il pourra utilement tenir compte des risques identifiés dans le cadre de la mise en œuvre de ses obligations au titre des articles 9 ou 55 du RIA ;
- l'analyse d'impact des systèmes d'IA à haut risque sur les droits fondamentaux prévue par l'article 27 du RIA pour certains dépoyeurs peut être pertinente pour l'analyse en phase de déploiement. En revanche, elle sera moins pertinente pour la phase de développement du modèle.

La CNIL poursuit ses travaux afin de préciser l'articulation entre les différents textes, notamment dans le cadre de l'élaboration des lignes directrices sur l'articulation entre le RIA et le RGPD actuellement en cours au sein du CEPD. La possibilité de combiner les différentes documentations a également été précisée par la CNIL dans ses [premières questions et réponses sur l'entrée en vigueur du RIA](#) et dans la [fiche IA n° 5 sur la réalisation de l'AIPD](#).

1. Sur les intérêts et bénéfices à prendre en compte

Synthèse des contributions

Plusieurs contributions **s'opposent à la distinction faite entre des intérêts de nature commerciale/non-commerciale** (comme la recherche scientifique ou l'intérêt public), au motif que ce ne serait pas pertinent pour déterminer l'atteinte susceptible d'être portée aux droits des personnes. Elles soulignent aussi que la distinction serait poreuse (la recherche publique peut être financée par le secteur privé et l'exploitation commerciale pourrait également bénéficier à un écosystème plus large que la seule entreprise).

Une autre contribution considère, à l'inverse, qu'il est **pertinent de distinguer l'intérêt de la communauté scientifique**, tout en insistant sur la nécessité de s'assurer que l'activité de recherche ne soit pas une façade pour la poursuite d'intérêts commerciaux.

Réponse de la CNIL

Un intérêt de nature purement commerciale, qui bénéficie uniquement à une entreprise privée, ne pourra pas avoir le même poids dans la mise en balance qu'un intérêt qui rejoint ceux plus larges de la société.

La CNIL rappelle toutefois que la notion de recherche scientifique au sens du RGPD n'exclut pas la recherche privée, pour autant qu'elle ne revête pas uniquement une finalité commerciale. Il sera nécessaire de démontrer que l'intérêt poursuivi relève de la recherche scientifique (voir pour plus de détails la fiche n° 2, qui précise les critères de définition de la recherche scientifique).

2. Sur les risques à prendre en compte

Synthèse des contributions

Une contribution considère que la structure de cette partie n'est pas claire et présente des incohérences.

Certaines contributions sont en faveur d'**une prise en compte extensive des risques lors de la phase de déploiement**, dont les risques d'atteinte à d'autres droits fondamentaux (ex. les risques de discrimination), les risques éthiques (dont certaines contributions considèrent qu'ils sont pris en compte de manière trop vague), les risques environnementaux, les risques qui pèsent sur les droits fondamentaux des personnes liés au cycle de vie de l'IA ou encore tous les risques qui résultent de l'impact du traitement lui-même, ainsi que de ses effets collatéraux.

À l'inverse, d'autres contributions considèrent qu'il faut **limiter les risques à ceux qui sont spécifiques à la phase de développement**, à l'exclusion de ceux liés à l'utilisation du modèle, au motif que cela sort du périmètre et que ces risques ne sont pas directement liés aux personnes concernées par le développement.

Réponse de la CNIL

Pour plus de clarté, la CNIL a simplifié la structure en faisant deux parties : les risques liés à au développement et les risques liés à l'utilisation du système d'IA.

Sur le périmètre des risques à prendre en compte lors de la phase de développement, les arguments avancés par les contributions n'ont pas été considérés comme remettant en cause l'équilibre proposé dans la fiche. En effet, il est essentiel de prendre en compte les risques systémiques qui apparaissent en phase de déploiement mais qui sont inhérents au développement du modèle (discrimination liée au manque de représentativité dans la base de données, manque de transparence, etc.), même s'il ne peut pas être exigé du développeur qu'il prenne en considération tous les risques qui peuvent se présenter lors du déploiement. Par ailleurs, cela rejoint la logique du mécanisme de gestion des risques prévu par l'article 9 du RIA concernant les systèmes d'IA à haut risque et l'avis sur les modèles d'IA du CEPD.

3. Sur les attentes raisonnables

Synthèse des contributions

De nombreuses contributions insistent sur le fait que **les personnes concernées ne peuvent pas raisonnablement s'attendre à ce que leurs données soient moissonnées ou collectées en ligne pour l'entraînement de systèmes d'IA**. Elles avancent pour cela les arguments suivants :

- il était impossible d'avoir conscience de la possibilité de tels traitements avant l'existence de ChatGPT en 2022 ;
- il n'y a pas de relation entre le responsable du traitement et la personne concernée ;
- les personnes ne sont pas informées des traitements ;
- si les personnes peuvent s'attendre à certains traitements des données rendues publiques (comme l'indexation d'une page web), elles ne peuvent pas raisonnablement s'attendre à ce que les données les concernant soient moissonnées pour former des jeux de données d'entraînement de systèmes d'IA.

À l'inverse, **une contribution considère que l'utilisation des données personnelles pour l'entraînement des modèles d'IA entre dans les attentes raisonnables des personnes**, notamment

depuis le lancement de ChatGPT en novembre 2022. Une contribution considère qu'il faut prendre en compte explicitement l'impact des évolutions technologiques en matière d'IA sur les attentes raisonnables des personnes.

Plusieurs contributions considèrent qu'il est nécessaire de prendre en compte **l'utilisation du système d'IA** (par ex. système d'hypertrucage) et **la source des données** (par exemple, selon le type de site concerné ou le paramétrage des publications moissonnées) pour évaluer les attentes raisonnables des personnes.

Réponse de la CNIL

L'analyse des attentes raisonnables des personnes doit s'effectuer au cas par cas, en tenant compte des circonstances spécifiques de chaque traitement. Pour préciser les cas qui pourraient entrer ou non dans les attentes raisonnables des personnes, la CNIL a :

- précisé dans le corps de la fiche les critères d'appréciation des attentes raisonnables des personnes, notamment au regard de ceux identifiés par l'avis 28/2024 du CEPD sur les modèles d'IA ;
- ajouté des exemples de traitements qui peuvent entrer ou non dans les attentes raisonnables.

4. Sur les mesures additionnelles

❖ Remarques générales

Synthèse des contributions

Plusieurs contributions s'interrogent sur le caractère obligatoire des mesures. Certaines contributions demandent une distinction entre ce qui relève des obligations, des recommandations et des bonnes pratiques pour faciliter la lecture et la compréhension. Certaines demandent d'ajouter un processus détaillé et des outils pratiques.

Certaines contributions considèrent que plusieurs des mesures citées créeraient des obligations qui iraient au-delà de ce qui est prévu par le RGPD ou par d'autres textes comme le RIA (comme le droit d'opposition discrétionnaire).

Réponse de la CNIL

La CNIL a légèrement restructuré cette partie pour la rendre plus pédagogique.

Sur le caractère prescriptif des mesures, la CNIL rappelle que **les mesures ne sont pas obligatoires en tant que telles mais constituent néanmoins des garanties fortes** dont la mise en œuvre pourra être nécessaire dans certains cas pour assurer la mise en balance. La pertinence et la nécessité de ces mesures devront être évaluées au cas par cas. Elles ne constituent en aucun cas des exigences supplémentaires à celles prévues par les textes.

❖ Sur les mesures recommandées

Synthèse des contributions

Certaines mesures ont été favorablement accueillies, comme les mesures additionnelles relatives à la transparence et à l'auditabilité des systèmes d'IA, le tatouage numérique, la mise en place d'un comité éthique, ou encore les mesures pour assurer la représentativité de la base de données.

Les contributions ont toutefois relevé des difficultés ou émis des réserves pour la mise en place de certaines mesures.

Certaines contributions considèrent que toutes les mesures citées ne sont pas nécessaires pour limiter l'atteinte portée aux droits des personnes, mais que les mesures suivantes seraient suffisantes, notamment en lien avec les risques de mémorisation et de régurgitation : exclusion des données indésirables des données moissonnées en ligne (notamment en évitant la collecte de sites contenant de nombreuses données personnelles), déduplication, techniques de post-entraînement (comme le RLHF pour apprendre par exemple au modèle à refuser de répondre à certaines requêtes), filtrage sur les sorties du modèle et à chaque entraînement.

Réponse de la CNIL

La CNIL considère que les mesures proposées dans le projet de fiche peuvent constituer des garanties fortes pour les personnes. Elle apporte, dans la fiche, des clarifications :

- Sur l'anonymisation/la pseudonymisation des données :

Ces mesures sont des garanties fortes qui ne s'opposent à l'utilisation des données personnelles, même dans des volumes importants, pour l'entraînement d'IA, dès lors qu'elles sont nécessaires au traitement (voir les fiches n°6 et 7 sur la minimisation). En outre, les jeux de données sont souvent hybrides : même s'il est difficile d'appliquer ces mesures à toute la base de données, il peut être pertinent d'avoir recours à des sous-jeux de données anonymisés ou synthétiques.

- Sur le recours aux données synthétiques :

Le recours aux données synthétiques peut être pertinent pour encourager l'émergence de techniques plus protectrices, d'autant plus que le développement des modèles par certains fournisseurs montre que c'est un champ de recherche potentiellement prometteur, malgré les limites actuelles.

- Sur la mesure relative à l'open source :

Dans l'attente du contenu dédié, qui doit préciser les conditions dans lesquelles la diffusion d'un modèle en source ouverte sera conforme à la réglementation en matière de protection de données, **la CNIL a supprimé cette mesure**. Celle-ci sera précisée dans les contenus ultérieurs.

- Sur la limitation par licence de la réutilisation :

Cette mesure est importante pour éviter l'utilisation des modèles à des fins malveillantes ou illicites. La CNIL a ajouté, dans cette fiche, une mesure visant à mettre en place des licences restreignant les usages pouvant permettre de réidentifier une personne.

- Sur les mesures pour éviter la régurgitation/mémorisation des données :

Pour préciser cette mesure, la CNIL liste plus précisément les mesures d'atténuation des risques qui pourront être adoptées, tout en renvoyant vers la fiche à venir sur le statut des modèles dans laquelle elles seront exposées de manière plus substantielles.

Les cas dans lesquels la mise en œuvre de ces mesures pourra permettre au responsable du traitement de conclure au caractère anonyme du modèle développé seront également précisés dans la fiche à venir sur le statut des modèles.

- Sur les informations supplémentaires en matière de transparence :

Afin de centraliser et préciser les bonnes pratiques recommandées en matière de transparence, la CNIL a inclus des mesures supplémentaires sur le contenu et les modalités de l'information fournie.

- Sur les mesures pour faciliter l'exercice des droits :

La mise en place d'un droit discrétionnaire peut constituer une garantie particulièrement forte pour permettre aux personnes de contrôler leurs données et ainsi être déterminante dans l'analyse de la mise en balance. La CNIL a par conséquent conservé cette mesure et apporté des précisions sur les modalités de sa mise en œuvre pour prendre en compte les éléments avancés par les contributions.

La fiche contient également une mesure relative à la mise en œuvre d'un droit à l'effacement des données pour assurer la cohérence avec les recommandations formulées par le CEPD dans son avis sur les modèles d'IA.

La CNIL a également précisé les mesures relatives à la facilitation de l'exercice des droits sur les données contenues dans la base de données ou dans le modèle, ainsi que sur la transmission des droits le long de la chaîne des acteurs, afin de préciser notamment l'articulation avec la fiche relative à l'exercice des droits (voir la fiche n° 10).

Fiche focus sur le moissonnage

A. Remarques générales

1. Sur la validité de l'intérêt légitime pour le moissonnage

Synthèse des contributions

Plusieurs contributions saluent favorablement l'approche de la CNIL pour cette fiche et la possibilité de mobiliser la base légale de l'intérêt légitime pour effectuer du moissonnage. Plusieurs d'entre elles soulignent, à cet égard, l'importance du moissonnage pour les développeurs d'IA en ce qu'il permet l'accès à des ressources en données diversifiées et dans des volumes importants. La possibilité d'avoir recours aux données accessibles en ligne est aussi perçue comme un enjeu de compétitivité, notamment avec les entreprises qui y ont de fait déjà eu recours jusqu'à présent.

À l'inverse, plusieurs contributions rappellent les risques importants liés au moissonnage et considèrent par principe que le moissonnage devrait être interdit ou, a minima, que la base légale de l'intérêt légitime ne pourrait pas être mobilisée.

Une série de contributions insistent en outre sur la question des attentes raisonnables. Si une minorité affirme que les personnes ont conscience que les données publiées en ligne sont susceptibles d'être exploitées, la plupart des contributions considèrent en revanche que le moissonnage ne pourrait pas entrer dans les attentes raisonnables des personnes.

Une contribution considère qu'il est surtout important de sensibiliser les personnes en amont de la publication de données en ligne.

Réponse de la CNIL

Les arguments avancés par les contributions n'ont pas conduit la CNIL à remettre en question son approche en matière de moissonnage : toutes les pratiques de moissonnage ne sont pas en soi illicites, mais et leur licéité dépendra d'une analyse au cas par cas, selon les modalités du traitement et les mesures mises en œuvre.

La CNIL précise dans la version finale de la fiche :

- les critères qui permettent de considérer que le moissonnage entre ou non dans les attentes raisonnables des personnes, en lien avec la fiche sur l'intérêt légitime, pour préciser les conditions dans lesquelles l'intérêt légitime pourra valablement être mobilisé ;
- des exemples de cas qui, sur la base de ces critères, peuvent valablement se fonder sur l'intérêt légitime ou non.

2. Sur l'articulation de cette fiche avec les autres textes ou positions

Synthèse des contributions

Plusieurs contributions s'interrogent sur l'articulation et la cohérence de cette fiche avec la doctrine antérieure de la CNIL, ainsi que celle de ses homologues.

Quelques contributions considèrent que cette fiche devrait également faire référence aux autres règles applicables en matière de moissonnage (notamment en matière de propriété intellectuelle, de réglementation applicable en matière de droits d'auteur et droits voisins) et de prendre en compte les risques d'atteinte qui en résultent pour l'évaluation de la licéité.

Réponse de la CNIL

La CNIL assure la cohérence de sa doctrine et participe activement aux travaux du CEPD sur les lignes directrices en cours d'élaboration en matière de moissonnage.

La CNIL a précisé que les questions en matière de propriété intellectuelle ne relèvent pas de sa compétence mais que les responsables de traitement doivent assurer la légalité des traitements au regard des autres réglementations en vigueur.

3. Sur le périmètre de la fiche

Synthèse des contributions

Plusieurs contributions s'interrogent sur le périmètre de la fiche, qui ne s'applique qu'aux développeurs de systèmes d'IA et non aux autres acteurs qui ont recours au *moissonnage*, ce qui créerait une **asymétrie dans les exigences qui leur sont applicables**. Elles considèrent que les recommandations devraient s'appliquer à toute pratique de *moissonnage*.

Réponse de la CNIL

Ces travaux portent sur la constitution de bases de données d'entraînement d'IA, qui présente des enjeux spécifiques, et non sur le moissonnage de manière générale, bien que ces recommandations puissent également être pertinentes pour le moissonnage à d'autres fins dans certains cas.

4. Les obligations des différents acteurs

Synthèse des contributions

Un certain nombre de contributions demandent des clarifications sur les obligations et responsabilités de chacun (hébergeur de site, organisme qui moissonne des données).

Réponse de la CNIL

La CNIL rappelle que ces obligations dépendent de la qualification du responsable du traitement au sens du RGPD (voir la [fiche n° 3](#) sur ce point).

En règle générale, si l'hébergeur d'un site dont les données sont moissonnées peut être dans certains cas responsable de leur diffusion ou mise à disposition, il n'est en revanche pas le responsable du traitement de leur réutilisation, c'est-à-dire en l'occurrence du moissonnage pour la constitution de la base de données. Il appartiendra au responsable du traitement (en général soit l'organisme qui effectue le moissonnage, soit celui pour le compte duquel le moissonnage est mis en œuvre) de s'assurer de la conformité du traitement. Pour plus d'informations sur les responsabilités des diffuseurs et réutilisateurs de données, [voir les fiches dédiées du guide sur l'ouverture et la réutilisation de données](#).

B. Sur la collecte de données sensibles

1. Sur la qualification de données sensibles

Synthèse des contributions

Quelques contributions considèrent, en se fondant sur la doctrine de la CNIL et du CEPD en matière de caméras et de caméras augmentées, qu'une donnée n'est considérée comme sensible que si les données sont utilisées pour déduire des données sensibles sur les personnes et propose d'appliquer cette doctrine au développement de l'IA, notamment dans la mesure où les fournisseurs ne traitent pas les données pour en déduire des informations sur les personnes prises individuellement.

Plusieurs contributions considèrent qu'il faudrait appliquer le même raisonnement pour les données hautement personnelles ainsi que les attributs de la personnalité (voix et image), voire considèrent que la voix et l'image devraient être considérées en elles-mêmes comme des données sensibles.

Réponse de la CNIL

Sur la définition des données sensibles, la CNIL rappelle que selon l'arrêt *Meta*, « cette interdiction de principe, prévue à l'article 9, paragraphe 1, du RGPD, est indépendante du point de savoir si l'information révélée par le traitement en cause est exacte ou non et si le responsable du traitement agit dans le but d'obtenir des informations relevant d'une des catégories particulières visées à cette disposition » (CJUE, 4 juillet 2023,

Meta, C-252/21, para. 69). La comparaison avec la doctrine du CEPD en matière de captation des images ne peut pas être étendue à tous les traitements compte tenu de sa spécificité.

Sur les spécificités du traitement de certaines catégories de données, la collecte de la voix et l'image à des fins de constitution d'une base de données d'entraînement ne constitue pas en tant que telles un traitement de données biométriques si le gabarit vocal ou du visage n'est pas utilisé à des fins d'identification unique de la personne. En revanche, en tant qu'attributs de la personne, la collecte et le traitement de ces données peuvent être particulièrement intrusifs pour les personnes. La CNIL a intégré un exemple qui porte sur la voix des personnes pour préciser ce point.

2. Sur la jurisprudence de la CJUE relative aux moteurs de recherche

Synthèse des contributions

❖ **Sur le principe de l'application de la décision de la Cour au-delà des moteurs de recherche**

Plusieurs contributions accueillent très favorablement l'approche de la partie sur les données sensibles en lien avec la jurisprudence de la CJUE (CJUE, 24 septembre 2019, GC e.a., C-136/17). Pour renforcer l'analogie, une contribution rappelle également le rôle des grands modèles de langage dans l'information des personnes.

Plusieurs contributions, **à l'inverse, remettent en question cette interprétation**, qui est inédite pour le moissonnage pour d'autres acteurs que les moteurs de recherche. Elles rappellent notamment que la CJUE avait mis en balance la responsabilité des moteurs de recherche au regard de leur rôle dans l'information du public, ce qui ne serait pas transposable au développement des modèles d'IA. Certaines soulignent également les divergences entre les autorités européennes sur l'interprétation de cet arrêt et appellent à une harmonisation à l'échelle européenne sur ce point.

❖ **Sur les mesures à mettre en œuvre pour supprimer les données sensibles**

Plusieurs contributions s'interrogent sur la suppression immédiate et automatique des données sensibles, au motif que cela irait au-delà de ce que prévoient les articles 6 et 9 du RGPD qui permettent le traitement de données sensibles dans certains cas.

Une contribution **considère, en outre, que ces mesures pourraient être techniquement irréalisables**, d'autant plus qu'il serait difficile d'interpréter ce qui peut constituer une donnée sensible au regard des moyens que cela requiert et de l'interprétation potentiellement extensive de ce qui constitue une donnée sensible.

Une association représentative de la société civile souligne qu'en pratique, les responsables de traitement ne sont pas en mesure d'empêcher la collecte de données sensibles lors du *moissonnage* ou du développement du modèle, parce qu'il n'est pas possible de faire la distinction entre les différentes catégories de données qu'ils détiennent (celles qui sont couvertes par le RGPD, celles pour lesquelles les personnes se sont opposées, les données sensibles, etc.). Tant que cette distinction n'est pas possible, cette contribution considère qu'il ne devrait pas être possible de permettre le traitement à grande échelle de données personnelles en ligne.

Réponse de la CNIL

Les arguments avancés n'ont pas conduit la CNIL à remettre en cause son interprétation de cet arrêt de la CJUE. Elle considère en effet que, dès lors que la configuration de l'outil ne cible pas les données sensibles, c'est-à-dire lorsque le moissonnage limite considérablement le risque de traiter des données sensibles, les obligations liées à leur traitement ne sont applicables que dans le cadre des responsabilités, des compétences et des possibilités du responsable du traitement. Elle insiste toutefois sur le fait qu'en l'absence d'une exemption au titre de l'article 9.2 du RGPD, le responsable du traitement est tenu de mettre en œuvre toutes les mesures permettant d'exclure la collecte de données sensibles en amont et dès qu'il vient à savoir qu'il traite des données sensibles. La CNIL participe activement aux travaux du CEPD pour l'harmonisation des positions des différentes autorités à l'échelle européenne.

Sur les mesures pour la suppression des données sensibles, la teneur des contributions semble témoigner d'une incompréhension de la doctrine de la CNIL. En effet, il s'agit d'exiger la suppression des

données sensibles non pertinentes dès que le responsable du traitement en a connaissance, sans préjudice toutefois de la possibilité de collecter ou conserver des données sensibles lorsqu'elles sont nécessaires pour la finalité poursuivie et que leur traitement peut se fonder sur une exception de l'article 9.2 du RGPD.

C. Sur les mesures de minimisation obligatoires

Synthèse des contributions

Certaines contributions **considèrent qu'il faudrait que la plupart des recommandations soient des mesures obligatoires** (comme exclure la collecte des sites qui s'opposent au moissonnage, l'exclusion de sites qui sont principalement constitués de données à caractère personnel, la diffusion de l'information sur la collecte et les droits, la mise en œuvre de mesures d'anonymisation et de pseudonymisation immédiatement après la collecte). Elles insistent sur la nécessité de limiter au maximum la collecte et le traitement des données sensibles.

À l'inverse, **plusieurs contributions remettent en question voire s'opposent assez nettement aux mesures obligatoires de minimisation indiquées** (en particulier en amont de la collecte).

S'agissant précisément de la définition des critères de collecte et l'application de filtres en amont, une contribution s'interroge sur la portée de l'obligation ainsi que sur l'existence de filtres en amont effectifs à ce jour. Plusieurs contributions considèrent que les seules exclusions qui pourraient être exigées en amont de la collecte seraient celles portant sur les contenus non-accessibles au public, les sites interdisant le moissonnage via des moyens techniques et automatisés (robots.txt, ai.txt), la mise en place d'une liste repoussoir centralisée.

S'agissant plus précisément de la suppression immédiate des données non-pertinentes, plusieurs contributions s'interrogent quant aux modalités précises de mise en œuvre de cette mesure, qui apparaît disproportionnée au regard du volume de données concernées. Elles soulignent également les difficultés à définir clairement ce qui serait une donnée non pertinente pour l'entraînement d'un modèle d'IA à usage général.

Réponse de la CNIL

La définition de critères précis de collecte et la suppression de données incidentes non nécessaires aux traitements sont des mesures minimales à mettre en œuvre. Compte tenu des difficultés techniques soulignées par les contributions, la CNIL considère que, dans les cas où il ne sera pas possible de filtrer certaines catégories de données (comme les données de mineurs ou des données sensibles), il sera nécessaire d'exclure de la collecte certains types de sites qui contiennent, par nature, ces catégories de données.

De plus, la CNIL ajoute à cette liste de mesures obligatoires le respect de l'utilisation de fichiers robots.txt et de CAPTCHA. En effet, ces mesures, qui peuvent également être mises en place par les sites pour protéger les droits d'auteur sur les contenus du site au titre de la directive sur les droits d'auteur et droits voisins, sont suffisamment répandues et expriment clairement l'opposition du site à l'utilisation des données qui y sont contenus.

D. Sur les mesures additionnelles

1. Sur l'approche de la partie

Synthèse des contributions

Plusieurs contributions saluent les garanties supplémentaires présentées, dont elles soulignent la nécessité pour garantir la mise en balance.

À l'inverse, plusieurs contributions remettent en cause l'approche de cette partie, au motif qu'elle manque de clarté sur la portée prescriptive de ces mesures, du risque d'incohérence qui résulterait de cette liste, et du manque de clarté sur l'articulation entre ces mesures et d'autres réglementations.

Réponse de la CNIL

Les mesures ne sont pas obligatoires en tant que telles, bien que certaines d'entre elles constituent des garanties fortes dont la mise en œuvre pourrait être nécessaire dans certains cas pour assurer la mise en balance. La pertinence et la nécessité de ces mesures devront être évaluées au cas par cas. Elles ne constituent pas des

exigences supplémentaires à celles prévues par les textes et n'entrent à ce titre pas en contradiction avec les autres réglementations.

2. Sur l'exclusion de certains sites

Synthèse des contributions

❖ Sur l'exclusion par défaut de la collecte de données à partir de certains sites

La plupart des contributions ont **accueilli cette mesure favorablement** en soulignant sa faisabilité et sa pertinence, en insistant sur les sites conçus comme des agrégateurs de données à caractère personnel, ceux qui contiennent des données sensibles, hébergent des contenus piratés, ou encore les sites pornographiques. Une contribution mentionne également qu'il faudrait exclure la collecte de données des sites avec un paywall ou une obligation de s'authentifier.

❖ Sur l'exclusion de la collecte sur les sites qui s'opposent clairement au moissonnage

Plusieurs contributions accueillent favorablement cette mesure et soulignent la **nécessité de respecter les CGU et licences de chaque site**, ce qui devrait impliquer pour les fournisseurs de mettre en place des mécanismes permettant de ne scraper que les sites dont les CGU autorisent l'utilisation des données par des sites tiers et à des fins tierces.

Plusieurs contributions mettent en avant certains procédés techniques pour exclure cette collecte.

Une contribution indique que, s'il est possible d'exclure les sites ayant été configurés pour repousser automatiquement le moissonnage, il ne serait pas matériellement possible d'exclure les sites pour lesquels cette opposition a été faite par un autre moyen (notamment quand c'est inscrit dans les CGU).

Réponse de la CNIL

Ces deux mesures sont très importantes, il apparaît nécessaire de les conserver *a minima* comme recommandations.

La CNIL précise l'articulation entre l'obligation, au titre du principe de minimisation, d'exclure certains sites lorsqu'ils contiennent des données dont le traitement n'est pas nécessaire et qu'il n'est pas possible de filtrer, et la mesure additionnelle consistant à exclure par défaut une liste prédéfinie de sites qui sont structurellement très intrusifs.

Par ailleurs, le respect des fichiers robots.txt comme une mesure obligatoire a été ajouté. La CNIL a en revanche conservé le respect des CGU comme recommandation ; elle rappelle que celles-ci devront être respectées lorsqu'elles sont opposables au responsable du traitement, au titre de leurs obligations contractuelles.

La CNIL a également inclus des recommandations adressées spécifiquement aux éditeurs de sites dans un encadré dédié, afin d'indiquer les mesures techniques existantes pour exprimer leur opposition au moissonnage, malgré l'absence de standardisation.

À cet égard, le projet actuel de code de bonnes pratiques pour les modèles d'IA à usage général de la Commission européenne encourage les organismes à soutenir les efforts de standardisation pour développer des standards interopérables permettant aux titulaires de droits de s'opposer au traitement de leurs données. La CNIL souhaite également encourager les responsables de traitement à participer à cet effort de standardisation, qui sera bénéfique pour l'exercice du droit d'opposition en matière de protection des données, et, si de nouveaux protocoles ou standards émergent, à s'y conformer.

3. Sur la limitation de la collecte aux données librement accessibles et manifestement rendues publiques

Synthèse des contributions

Quelques contributions **accueillent favorablement** cette mesure.

Plusieurs contributions **s'interrogent** sur ce qu'il faut entendre par « données librement accessibles », notamment au regard des différentes restrictions d'accès aux sites. Sur la distinction entre les données accessibles seulement avec un compte et sans compte, une contribution considère que seul le moissonnage des dernières devrait être possible.

Deux contributions **considèrent que limiter la collecte de données à celles qui ont été manifestement rendues publiques par les personnes concernées est techniquement irréalisable**, compte tenu de la nature non structurée des données et de l'absence d'identification des personnes. Une contribution estime également que cette mesure va au-delà de ce qu'exige le RGPD car la collecte de données manifestement rendues publiques est l'une des exceptions prévues pour la collecte de données sensibles et non pour la collecte de données personnelles de manière générale.

Réponse de la CNIL

La CNIL considère que cette recommandation peut être une mesure forte pour s'assurer que les données publiées avaient vocation à être partagées au plus grand nombre.

La notion de « manifestement rendues publiques » pourrait en revanche être trop restrictive, car elle renvoie aux critères de la jurisprudence concernant le traitement de données sensibles et nécessiterait d'analyser l'intention des utilisateurs, au-delà du simple paramétrage. La CNIL a par conséquent modifié la formulation de cette mesure pour recommander de limiter la collecte aux données librement accessibles, en précisant que cette notion fait référence aux contenus accessibles à tout utilisateur non inscrit sur le site en question et sans création d'un compte, et que les personnes doivent avoir conscience de rendre leurs données publiquement accessibles.

4. Sur la diffusion des informations relatives à la collecte et aux droits des personnes

Synthèse des contributions

Si quelques contributions sont favorables à cette mesure sur la transparence, plusieurs d'entre elles sont revenues sur leur opposition à la publication des sites concernés par le moissonnage (atteinte au secret des affaires, sécurité des modèles, ni prévu par le RGPD ni le RIA, etc.).

D'autres mesures de transparence sont suggérées, comme des informations sur la manière dont les IA sont entraînées, des informations spécifiques sur les sources de données et les pratiques de collecte, ainsi que des informations sur les droits des individus, qu'il faudrait accorder avec le code de bonnes pratiques pour les modèles d'IA à usage général de la Commission européenne.

Une contribution considère que cette recommandation en matière d'information devrait **s'appliquer à l'hébergeur des sites** contenant des données susceptibles d'être moissonnées à et non au responsable du traitement qui moissonne les données.

Réponse de la CNIL

Le projet de fiche recommandait de diffuser le plus largement possible les informations relatives à la collecte et aux droits des personnes (par exemple par l'intermédiaire d'articles en ligne), en s'assurant de publier une liste mise à jour des sites concernés par les pratiques de moissonnage.

Sur le contenu de l'information concernant les sources des données, la CNIL précise que cette recommandation est en ligne avec celle qui a été formulée dans le cadre de la [fiche sur l'information](#), à laquelle la CNIL renvoie explicitement.

Sur le lieu de l'information, il serait en effet pertinent que l'information soit fournie aux personnes sur le site dont les données sont moissonnées. Si cette obligation ne peut pas incomber à l'hébergeur des sites, la CNIL a inclus une bonne pratique similaire à ce qui est préconisé dans le guide sur l'ouverture et la réutilisation des données, en invitant les développeurs à se rapprocher des sites où sont hébergées les données pour multiplier les lieux d'information au plus proche des personnes dont les données sont collectées.

5. Sur la possibilité pour les personnes de s'opposer au traitement

Synthèse des contributions

❖ Sur la mise en place d'une liste repoussoir

Quelques contributions accueillent favorablement cette mesure. Certaines contributions demandent des précisions sur le fonctionnement d'une telle liste repoussoir dans le contexte du moissonnage.

De nombreuses contributions s'opposent à l'inverse à cette mesure, notamment en raison de son infaisabilité (absence de lien des fournisseurs avec les personnes concernées, difficulté pour identifier les personnes concernées et retracer la source des données, ce qui empêche de permettre aux individus d'exercer leur droit d'opposition avant le moissonnage, etc.).

Plusieurs contributions considèrent qu'une telle liste repoussoir **devrait être mise en place à l'échelle nationale, voire européenne**, par un texte législatif ou réglementaire, et pilotée par un organisme public, sur le modèle de Bloctel. Une contribution s'interroge sur la possibilité pour la CNIL d'être le référent pour une telle liste.

Une contribution considère que **d'autres mesures sont plus pertinentes et efficaces** dans le cadre du développement des modèles d'IA, comme la possibilité pour les personnes de demander la suppression de leurs données personnelles des données de sortie du modèle.

❖ Sur le droit d'opposition discrétionnaire

Une contribution **salue la mise en place d'un droit d'opposition discrétionnaire** et considère que les entreprises qui font du moissonnage et les développeurs devraient envisager de fournir aux utilisateurs la possibilité de s'opposer au moissonnage des données publiquement accessibles de manière standardisée.

Plusieurs contributions **s'interrogent sur les modalités de mise en œuvre** de cette recommandation et soulignent les difficultés techniques et organisationnelles que cela présenterait. Plusieurs contributions s'interrogent quant à **l'acteur qui en aurait la responsabilité** (site, plateforme, fournisseur), notamment au motif qu'il ne serait possible de s'opposer au traitement que sur la plateforme mettant en ligne les données.

Plusieurs contributions **s'opposent plus fermement** à cette mesure, au motif que la mise en place d'un droit d'opposition discrétionnaire va plus loin, voire, serait contraire à l'article 21 du RGPD, qui aurait déjà opéré un équilibre des intérêts et libertés en cause.

Réponse de la CNIL

De manière générale, en cas de moissonnage de données accessibles en ligne, la CNIL encourage le développement de solutions techniques qui permettraient de faciliter le respect de l'exercice du droit d'opposition en amont de la collecte des données. Il existe, pour certains traitements, des mécanismes de « liste repoussoir », qui pourraient être transposés lorsque c'est adapté au regard du traitement de données mis en œuvre. La CNIL précise ces éléments dans le corps de la fiche, en lien avec les recommandations formulées dans [la fiche n° 10 sur l'exercice des droits](#).

6. Sur le projet de registre des organisations traitant des données collectées par moissonnage à des fins de développement de systèmes d'IA

Synthèse des contributions

Les contributions ayant répondu sur ce point étaient partagées mais une grande partie des acteurs susceptibles de pouvoir y figurer s'y est opposée, parfois assez nettement.

Les principaux arguments invoqués contre le registre étaient notamment les suivants :

- La probabilité que les entreprises n'y ait pas recours compte tenu du caractère facultatif de l'inscription et absence de garantie en termes de conformité, ce qui risquerait de nuire à l'utilité du projet ;
- Le risque de fragmentation à l'échelle européenne, compte tenu du caractère national de l'initiative ;
- La limitation du périmètre à l'IA, qui créerait d'après certains acteurs une inégalité avec les acteurs qui font de moissonnage de données pour d'autres finalités ;
- Les difficultés pour garantir l'exactitude des informations contenues dans le registre.

Plusieurs contributions ont suggéré la mise en place d'un cadre normatif dédié à tel registre par les pouvoirs publics, à l'échelle française ou européenne, ce qui permettrait la mise en place d'une solution harmonisée.

Certaines contributions ont également émis des suggestions alternatives sur la forme du registre et les informations qu'il contiendrait.

Réponse de la CNIL

Compte tenu des retours des différents acteurs et du risque d'une faible adhésion de l'écosystème, qui risquerait de priver le dispositif de son utilité pour les personnes concernées, la CNIL a pour le moment décidé de suspendre le projet de registre, malgré l'intérêt qu'il pouvait présenter en termes de transparence.