Recommandations

Réutilisateurs de données publiées sur Internet

Publié le 12/06/2024

CNIL.
COMMISSION NATIONALE
INFORMATIQUE & LIBERTÉS

Introduction

À qui s'adresse ces recommandations ?

Ces recommandations s'adressent aux réutilisateurs de tous types de données personnelles publiées sur Internet – données mises à disposition à des fins de réutilisation (*open data*) et autres données librement accessibles en ligne (ex. : sites d'information et blogs, sites commerciaux, informations partagées par des internautes sur les réseaux sociaux, ...) – **par des personnes physiques ou morale, publiques ou privées, collectant les données en vue d'une exploitation de celles-ci pour leur propre compte**.

Leur contenu pourra également intéresser les personnes concernées par les traitements mis en œuvre par ces réutilisateurs : elles y trouveront, en particulier, des informations sur les droits dont elles disposent pour conserver la maîtrise des usages qui sont faits de leurs données.

Quelle est leur vocation ?

Ces recommandations complètent, tout en élargissant le champ d'étude, le guide co-édité en 2019 avec la CADA sur l'ouverture et la réutilisation des données publiques. Par ailleurs, elles se trouvent prolongées par des fiches « cas d'usage ».

Ces « fiches principes », à caractère pratique et opérationnel, fournissent une grille d'analyse générale permettant à tout réutilisateur de données en ligne de cheminer le plus rapidement et efficacement possible sur les questions « informatique et libertés » structurantes et éléments de réponse pertinents.

À noter

Ces fiches, adoptées à la suite d'une consultation publique, constituent un cadre qui permet d'accompagner les organismes dans leur mise en conformité. Elles rappellent les obligations posées par la règlementation et formulent des recommandations pour s'y conformer. **Ces recommandations ne sont pas contraignantes : les responsables de traitement peuvent s'en écarter, à condition de pouvoir justifier leur choix et sous leur responsabilité**. Certaines recommandations sont également formulées à titre de bonnes pratiques et permettent d'aller plus loin que le respect de la règlementation.

Table des matières

Fiche n°1 : Quelle qualification juridique pour les réutilisateurs de données ?

Toute personne physique ou organisme traitant des données personnelles doit au préalable s'interroger sur sa qualification au sens du RGPD, qualification dont vont dépendre ses obligations. Pour un traitement donné, il est possible d'être responsable de traitement, sous-traitant ou responsable de traitement conjoint. Il incombe aux acteurs de déterminer leur qualification au cas par cas. La présente fiche rappelle les définitions de ces différentes notions et explique comment les appliquer en cas de réutilisation de données publiquement accessibles sur Internet.

Le responsable de traitement

Le responsable de traitement est la personne physique ou morale, l'autorité publique, le service ou un autre organisme qui **détermine les finalités et les moyens du traitement, c'est-à-dire qui décide du « pourquoi » et du « comment » de l'utilisation de données personnelles.**

Focus

Qu'est-ce qu'une finalité ?

La finalité est le but précis pour lequel les données sont traitées. Elle doit être déterminée, explicite et légitime.

Il peut s'agir, par exemple, de :

- la réutilisation de données publiquement accessibles à des fins de constitution et de diffusion d'un annuaire de professionnels[1] ;
- la réutilisation de données publiquement accessibles à des fins de prospection commerciale[2].

En présence d'un traitement impliquant une réutilisation de données publiées sur Internet, une analyse au cas par cas est toujours nécessaire pour déterminer qui est responsable de traitement.

Lorsqu'une personne décide d'exploiter des données et détermine la finalité (l'objet, le but de la réutilisation) et la caractéristique principale du traitement (utiliser des données publiées sur Internet), elle est en principe responsable de traitement.

Exemple

Ce serait par exemple le cas **de la personne qui décide de constituer une base de données à partir de données publiquement accessibles pour l'exploiter commercialement auprès de différents clients.**

Cette personne et chacun de ses clients pourraient être qualifiés de responsables de traitements distincts (la première pour les opérations de recueil et de commercialisation des données, les seconds pour les utilisations qu'ils en font), à condition que chaque traitement en cause puisse être séparé l'un de l'autre (et que leurs finalités et moyens ne soient pas déterminés conjointement comme cela est détaillé ci-dessous).

Ce principe est valable, y compris lorsque la personne à l'initiative du traitement de données personnelles décide de recourir, pour sa mise en œuvre, aux outils ou services d'un prestataire.

[1] Voir la fiche pratique dédiée à ce cas d'usage.
[2] Voir la fiche pratique dédiée à ce cas d'usage.

Focus

Les contrats d'adhésion

Le fait de recourir à un logiciel de traitement de données personnelles conçu par un autre acteur, et sur lequel son utilisateur ne peut qu'effectuer certains paramétrages (voire aucun), ne dispense pas ce dernier de sa qualité de responsable de traitement, dès lors que c'est lui qui a décidé d'utiliser telles et telles données personnelles avec ce logiciel, pour une certaine finalité (CJUE, deuxième chambre, 29 juillet 2019, Fashion ID GmbH & Co. KG, C-40/17).

Un tiers peut donc déterminer des moyens de traitement indépendamment du traitement opérationnel (ex. : matériel ou logiciel utilisé pour le traitement, méthode de stockage des données, etc.), ou fournir un service « clé en main », sans pour autant être responsable de traitement. Si ce tiers est ensuite conduit, sur instruction de son client, à manipuler lui-même les données personnelles pour le traitement, il agira en tant que sous-traitant. Dans certains cas, il pourra être regardé comme responsable conjoint du traitement et non simple sous-traitant, notamment si la mise en œuvre du traitement qu'il opère dépasse la seule prestation de service au bénéfice de ses clients (voir les points suivants relatifs à ces qualifications).

Le responsable conjoint du traitement

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

Exemple

Ce serait par exemple le cas d'organismes poursuivant une finalité commune, comme des établissements universitaires qui décideraient de mener en partenariat une recherche sur un sujet particulier à partir du recueil et de l'analyse de données collectées sur internet.

Lorsque plusieurs acteurs traitent des mêmes données pour des finalités propres, la question de savoir s'ils doivent être considérés comme des responsables de traitements distincts ou comme des responsables de traitement conjoints peut être délicate à opérer. Il convient alors de déterminer, pour le traitement en cause, si chacun des acteurs a ou non exercé une influence déterminante sur sa ou ses finalité(s), ainsi que sur ses conditions de mise en œuvre.

Focus

Les licences de réutilisation

Le fait que le diffuseur des données ait encadré les réutilisations futures par le biais d'une licence ou de conditions générales d'utilisation, notamment pour assurer aux personnes concernées une certaine prévisibilité sur l'usage de leurs données, ne le rendra pas nécessairement responsable de ces réutilisations.

En effet, d'une part, l'autorisation ou l'interdiction par le diffuseur de certaines catégories de réutilisation ne permet pas de considérer que la finalité de la réutilisation a été conjointement déterminée par ce dernier si le réutilisateur conserve un degré d'influence autonome sur ses propres traitements, à commencer par la décision d'utiliser ou non les données.

D'autre part, déterminer conjointement la finalité n'est pas suffisant pour qualifier une responsabilité conjointe si les moyens du traitement sont définis distinctement.

- Exemple de cas où la responsabilité conjointe des parties prenantes à un même traitement pourrait être retenue, en l'absence d'une décision commune initiale mais du fait de décisions convergentes, complémentaires et nécessaires à sa mise en œuvre : traitement réalisé, à la demande d'un employeur, par un chasseur de tête qui collecterait notamment des données sur internet (candidats potentiels à des offres d'emplois) à la fois pour répondre aux besoins précis de son client (identifier des candidats adéquats) et pour alimenter sa propre base de données, qu'il a constituée et utilise pour l'exécution de tous les contrats qu'il signe avec ses clients.

En cas de responsabilité conjointe, chaque responsable de traitement doit s'assurer de la licéité du traitement, notamment en définissant, dans le cadre d'un « accord » et de manière opérationnelle et transparente, leurs obligations respectives (article 26 du RGPD).

La forme de cet accord n'est pas précisée par le RGPD. L'essentiel est que les parties s'engagent mutuellement, via un contrat par exemple, sur « qui fait quoi » pour que soient respectées toutes les règles relatives à la protection des données personnelles. Ainsi, cet accord a notamment vocation à préciser les modalités d'exercice et de prise en compte des droits « informatique et libertés ».

À cet égard, il est à noter que si les demandes des personnes concernées ont vocation à être prises en charge par le ou les responsables qui y est/sont désigné(s), elles gardent la liberté d'exercer leurs droits auprès de n'importe lequel des responsables de traitement conjoints.

Le sous-traitant

Le sous-traitant **agit pour le seul compte du (ou des) responsable(s) de traitement**. Il ne peut traiter les données pour son propre compte (sauf exception), et ne doit les traiter que sur instructions documentées du responsable de traitement.

Exemples de sous-traitants

- Un éditeur et gestionnaire d'une interface de programmation applicative (« API ») de moissonnage de données permettant à des réutilisateurs de récupérer de manière automatisée des données en ligne, dès lors qu'il ne détermine pas la finalité et les moyens de chaque utilisation (tels que les données collectées) sollicitée par ses différents clients.
- Un prestataire de services qui réaliserait, sur demande de ses clients et à partir de paramètres prédéfinis ou validés par ces derniers, des opérations de collecte et d'analyse de données publiquement accessibles.

En situation de sous-traitance, un contrat doit être conclu entre le sous-traitant et le responsable de traitement. Ce contrat doit contenir toutes les mentions prévues par l'article 28 du RGPD. Par ailleurs, le responsable de traitement doit s'assurer que son sous-traitant présente des garanties suffisantes pour la conformité des traitements

Pour approfondir :

- Le guide du sous-traitant (PDF, 583 ko), cnil.fr

Fiche n°2 : Comment identifier la base légale de son traitement ?

La base légale d'un traitement est ce qui donne le droit à un organisme de traiter des données personnelles. L'identification d'une base légale est donc une première étape indispensable pour assurer la conformité d'un de réutilisation de données publiquement accessibles sur Internet.

L'obligation d'identifier une base légale

Pour être licite, tout traitement doit se fonder sur l'une des six « bases légales » (ou « fondements juridiques ») prévues par le RGPD.

Une base légale valable doit ainsi être identifiée en amont de la mise en œuvre du traitement, au cas par cas, de manière adaptée à la situation et au type de traitement en cause (statut de son responsable, objectifs poursuivis, obligations réglementaires, enjeux pour les personnes concernées, etc.).

Attention

Le fait que des données soient « publiquement accessibles », en étant notamment **en libre accès sur Internet, ne signifie pas qu'il est possible de se passer d'une base légale ou de réutiliser ces données sans conditions.**

En matière de réutilisation de données publiquement accessibles, les principales bases légales envisageables[3] sont :

- l'obligation légale : le traitement est imposé à l'organisme, public ou privé, par des dispositions législatives ou réglementaires ;
- la mission d'intérêt public : le traitement est nécessaire à l'exécution d'une mission d'intérêt public telle que définie par des dispositions légales ;
- l'intérêt légitime : le traitement est nécessaire à la poursuite d'intérêts légitimes de l'organisme qui traite les données ou d'un tiers, sous réserve que ne prévalent pas les droits et intérêts des personnes dont les données sont traitées ;
- le consentement : la personne a consenti au traitement de ses données, de manière libre, spécifique, éclairée et univoque.

L'identification d'une base légale appropriée est d'autant plus importante qu'elle a aussi des conséquences sur les droits des personnes concernées.

Pour approfondir

- Les bases légales, cnil.fr

[3] Aucune des bases légales prévues à l'article 6 du RGPD ne doit être proscrite par principe.

Questions à se poser pour identifier sa base légale

À noter

Si les conditions propres à la base légale envisagée par chacune des questions ci-dessous ne sont pas remplies, l'organisme doit, soit modifier les paramètres de son projet de traitement pour parvenir à les respecter, soit rechercher une autre base légale.

1. Cette réutilisation de données personnelles est-elle nécessaire au respect d'une obligation légale ?

Pour que la base légale « obligation légale » puisse être retenue, il faut :

- qu'un texte suffisamment clair et précis exige de l'organisme, la mise en œuvre du traitement dans les conditions projetées ;
- qu'il n'existe, en d'autres termes, pas de moyen moins intrusif permettant de respecter les dispositions légales en cause.

Illustration

La CNIL a reconnu la validité de cette base légale pour la recherche par les assureurs des bénéficiaires de contrats d'assurance-vie non réclamés, notamment sur le web ou sur des réseaux sociaux publiquement accessibles, à condition que cette recherche ne soit pas automatisée et ne conduise pas à une collecte massive[4].

Néanmoins, l'obligation légale pourra rarement être invoquée en pratique.

2. Si non, est-elle soumise obligatoirement au consentement des personnes concernées en vertu d'une disposition légale ?

Certains textes imposent de fonder le traitement sur la base légale du consentement. Il en va ainsi, sauf exceptions (p. ex. : prospection « de professionnels à professionnels »), des traitements ayant pour objet la prospection commerciale par voie électronique (art. L34-5 du code des postes et communications électroniques), ou encore des traitements emportant, avec ou son profilage préalable, des prises de décisions automatisées aux effets significatifs pour les personnes (art. 22 du RGPD).

3. Si non, est-elle nécessaire à l'exercice d'une mission d'intérêt public qui m'a été légalement confiée ?

La possibilité de se fonder sur la base légale de la « mission d'intérêt public » suppose :

- que la mission dans laquelle s'inscrit le traitement soit prévue par un texte normatif applicable au réutilisateur ;
- que la réutilisation des données permette d'exercer *spécifiquement* cette mission (pas le cas si elle vise un objectif sans rapport particulier avec celle-ci ou trop éloigné de ses particularités), de manière pertinente et appropriée.

S'il ne s'agit pas d'une obligation, plus ces dispositions sont précises, plus il est facile pour les organismes concernés de recourir à cette base légale.

[4] Il s'agit de respecter la loi n° 2014-617 du 13 juin 2014 relative aux comptes bancaires inactifs et aux contrats d'assurance vie en déshérence.

Illustration

C'est le cas du pôle d'expertise de la régulation numérique (PEReN) qui dispose d'un pouvoir de réutilisation des données publiquement accessibles sur les sites des opérateurs de plateforme afin de réaliser des expérimentations ayant notamment pour objet de concevoir ou évaluer des outils techniques destinés à la régulation des opérateurs de plateformes en ligne[5].

4. Si non, la réutilisation de données peut-elle se fonder sur la poursuite d'un intérêt légitime ?

La réutilisation doit répondre à un intérêt légitime du réutilisateur et ne pas porter une atteinte excessive aux droits et intérêts des personnes concernées par les données. Les bénéfices attendus de la réutilisation, pour le responsable de traitement ou pour les tiers, les **attentes raisonnables** des personnes concernées par le projet de traitement, **ainsi que les impacts/risques qui en résultent et les garanties envisageables pour les réduire**, sont les critères déterminants pour apprécier si l'intérêt légitime du réutilisateur peut valablement justifier la mise en œuvre du projet.

Trois cas de figure doivent être distingués :

- **Premier cas :** les données mises à disposition du public le sont en vertu de la législation sur l'*open data*, dont le but est précisément de permettre leur libre réutilisation en raison de l'intérêt qu'elles sont susceptibles de présenter pour des tiers. Dans ce cas, si la finalité poursuivie par le réutilisateur est elle-même légale, le traitement reposera presque toujours sur un intérêt légitime. On ne peut exclure que, malgré la mise à disposition des données pour réutilisation, certaines formes de traitement particulièrement intrusives portent aux intérêts des personnes concernées une atteinte excessive, qui les rendrait illégales sans le consentement de celles-ci.
- **Deuxième cas :** la personne qui a diffusé la base de données l'a fait volontairement, pour permettre la réutilisation, mais celle-ci n'est pas prévue par le cadre légal. Dans ce cas, l'intérêt légitime du réutilisateur est le plus souvent acquis si la diffusion de la base était elle-même légale. Il convient en effet de tenir compte de ce que la mise en ligne de la base de données n'est elle-même légale que si le diffuseur a pu estimer que cette publication, qui expose les données à toutes formes de réutilisation, ne porte pas une atteinte disproportionnée aux personnes, ou, à défaut, a recueilli le consentement des personnes concernées (en tout état de cause, les personnes auront dû, en principe, être préalablement informées de la diffusion). Malgré cela, il convient de s'interroger à chaque fois, pour vérifier si l'utilisation qui est faite de la base peut se fonder sur un intérêt légitime ou nécessite le consentement des personnes.
- **Troisième cas :** des données personnelles ont été publiées sur internet mais ne l'ont pas été dans une perspective d'*open data*, mais simplement dans le cadre de l'activité du responsable de traitement (données figurant sur les réseaux sociaux, données d'articles de presse, données présentes sur un site de petites annonces etc.). Si le caractère d'ores et déjà public de la donnée doit être pris en compte pour apprécier l'atteinte que leur réutilisation pourrait porter à la vie privée des personnes concernées, et les risques auxquels elles pourraient être exposés, l'intérêt légitime du réutilisateur doit être apprécié au cas par cas.

Dans le cadre de la mise en balance par le réutilisateur des intérêts en présence, plusieurs éléments peuvent être pris en compte, en particulier :

- le contexte de la collecte des données, leur finalité initiale, et le contexte de la publication des données en ligne (p. ex. : publication prévue ou non par un cadre légal, intervenant sur un site d'échanges d'ordre personnel ou professionnel) ;
- les catégories de personnes (p. ex. : personnalités publiques, mineurs) et de données concernées, plus ou moins sensibles (p. ex. : données professionnelles ou données relatives à la vie privée) ;
- les éventuelles restrictions découlant d'une licence de réutilisation ou des conditions générales d'utilisation d'un site web (p. ex : interdiction de procéder à des moissonnages de données en ligne à des fins de prospection commerciale) ;
- les mesures susceptibles d'être adoptées pour limiter l'impact du traitement sur les personnes (ex. : anonymisation à bref délai, pseudonymisation, droit d'opposition inconditionnel).

[5] Conformément au décret n° 2022-603 du 21 avril 2022 fixant la liste des autorités administratives et publiques indépendantes pouvant recourir à l'appui du pôle d'expertise de la régulation numérique et relatif aux méthodes de collecte de données mises en œuvre par ce service dans le cadre de ses activités d'expérimentation.

Illustrations

Par exemple, « l'intérêt légitime » pourrait valablement fonder :

- la rediffusion, par des sociétés spécialisées dans l'information légale et financière, des données d'entreprises diffusées par les administrations chargées de leur mise à disposition au public (INPI et INSEE) ;
- l'exploitation, à des fins de réalisation d'études sur l'évolution des prix de l'immobilier, de la base de données « demandes de valeurs foncières / DVF » obligatoirement diffusée par la DGFiP et faisant état des transactions immobilières des cinq dernières années, sans identification directe des personnes concernées) ;
- la collecte par un employeur de données publiées sur un réseau social professionnel afin de contacter un candidat potentiel, dès lors que les personnes concernées, en les partageant sur un tel site s'attendent raisonnablement à ce type de réutilisations (voir la fiche n°14 du guide de la CNIL sur le recrutement) ;
- l'exploitation, par l'établissement de statistiques intégrant des mécanismes de pseudonymisation et d'anonymisation, de données publiées sur des blogs, forums, réseaux sociaux, etc., à des fins de mesure de l'opinion sur un sujet, une personne ou un produit spécifique (détection de tendances, corrélations) ;
- la réutilisation, dans le strict respect des droits des personnes concernées (droit de rectification, en particulier), des données publiées sur des sites institutionnels aux fins d'améliorer la transparence des conditions d'exécution par les élus de leur mandat politique (élaboration et publication de contenus synthétisant leurs activités : taux de présence, nombre de rapports, sens de leurs votes sur différentes thématiques, etc.).

À l'inverse, « l'intérêt légitime » n'apparaît pas pouvoir valablement fonder :

- l'exploitation des coordonnées publiques des DPD / DPO, diffusées en *open data* par la CNIL (art. 83 du décret d'application de la loi Informatique et Libertés), à des fins de communication sur des sujets sans lien avec leur activité professionnelle (p. ex. : prospection politique) ;
- la réutilisation des coordonnées électroniques publiées par des utilisateurs de sites de petites annonces entre particuliers à des fins de commercialisation de bases de prospection (p. ex. : auprès d'agences immobilières, s'agissant d'annonces dédiées à la location / vente de logements) ;
- le recours à un procédé particulièrement intrusif et massif qui permet de récupérer les images présentes de façon éparse sur Internet de millions d'internautes en France, afin d'alimenter un système de reconnaissance faciale pouvant être utilisé par des États à des fins policières (voir à ce sujet la sanction publique rendue par la CNIL, n° SAN-2022-019 du 17 octobre 2022) ;
- le traitement mis en œuvre par une entreprise éditrice d'un site web compilant de multiples informations sur d'innombrables personnes (photos, vidéos, posts, CV, liens vers les pages web où elles apparaissent, etc.), tirées de diverses sources ouvertes (réseaux sociaux, blogs, etc.), et les restituant aux internautes sous forme de profils à partir de la saisie de données nominatives dans un moteur de recherche.

À noter : tous les cas de rediffusion de données déjà publiées ne pourront pas se fonder sur la base légale de l'intérêt légitime. Il existe effectivement une différence d'enjeux et d'impacts entre l'établissement de profils nominatifs de particuliers, sur la base d'une agrégation de toutes les données en ligne les concernant, et la « simple » rediffusion de données d'entreprises par ailleurs diffusées par des administrations dans le cadre de leurs missions légales.

Pour ce type de traitements, auxquels les personnes concernées ne sauraient raisonnablement s'attendre, seul le recueil de consentements valables auprès de celles-ci devrait permettre de garantir leur licéité.

5. Si non, puis-je recueillir un consentement valable auprès des personnes concernées ?

Pour être valable, il doit respecter un certain nombre de conditions.

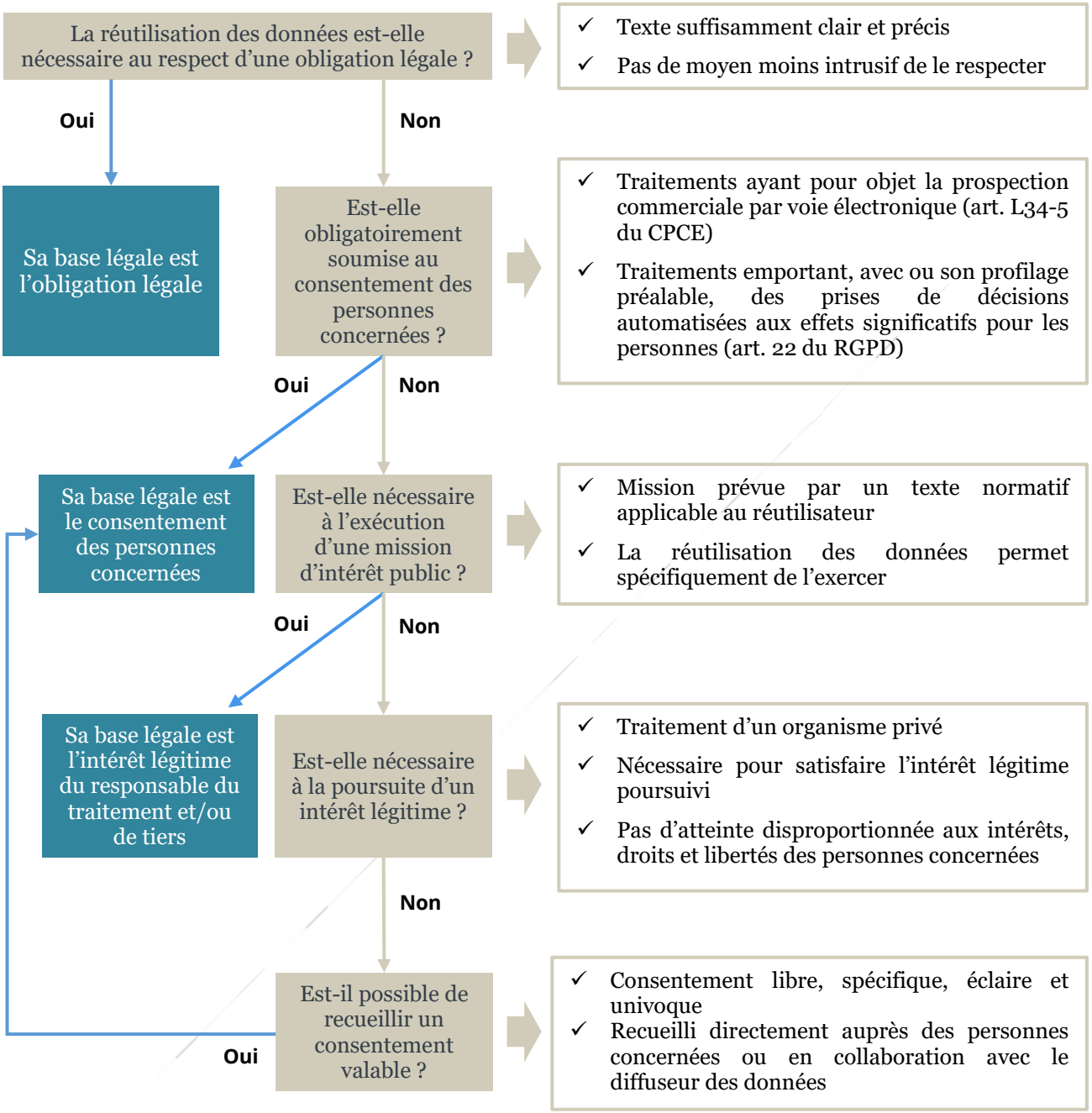
S'agissant des modalités de recueil du consentement, les possibilités suivantes méritent d'être évoquées.

- Dans l'hypothèse où l'organisme souhaitant réutiliser les données dispose du moyen de contacter les personnes concernées (p. ex. : utilisateurs de réseaux sociaux ayant un profil public comportant leurs adresses mail ou un outil de contact accessible par toute personne, blogueurs mettant à disposition un formulaire de contact), il pourra, sans difficultés, solliciter directement auprès d'elles le recueil de leur accord.

- Dans l'hypothèse inverse, l'organisme souhaitant réutiliser les données peut demander à celui les ayant initialement publiées d'informer lui-même les personnes concernées de son projet de traitement et d'inviter celles n'y étant pas opposées à se rapprocher du réutilisateur pour lui fournir leur consentement. Cela suppose toutefois que l'organisme ayant publié les données soit en lien avec les personnes concernées et accepte de jouer le rôle d'intermédiaire (notamment, compte tenu du fait qu'il ne serait pas légitime à communiquer directement les données de contact qu'il détient au réutilisateur).

- Il existe enfin des cas où le consentement d'une personne à une réutilisation de ses données pour certaines finalités pourrait être donné par celle-ci au moment de la mise en ligne des informations la concernant, par exemple au moyen d'une case à cocher lors de la création de son compte sur un réseau social ou sur un site de petites annonces. Une attention particulière doit alors être portée, d'une part à la liberté du consentement, qui doit souvent pouvoir être accordé ou refusé finalité par finalité (le consentement à la publication n'emporte pas un consentement à toute forme de réutilisation) et, d'autre part, aux informations devant être portées à la connaissance des personnes, qui incluent notamment l'identité de chaque responsable de traitement réutilisateur. Cette option implique elle aussi une collaboration entre le diffuseur et le réutilisateur des données.

Schéma récapitulatif : identifier la base légale du traitement



Fiche n°3 : Comment informer les personnes concernées ?

Le RGPD impose d'informer les personnes concernées de la réutilisation de leurs données accessibles sur internet. La CNIL fait le point sur les mesures à prendre pour respecter cette obligation et précise les cas dans lesquels une information publique générale pourra suffire.

Pourquoi assurer la transparence des traitements ?

Le principe de transparence oblige les organismes collectant des données personnelles à en **informer les personnes afin qu'elles comprennent les usages** qui seront faits de leurs informations (pourquoi, comment) **et soient en mesure d'exercer leurs droits** (opposition, accès, rectification, etc.). Il contribue ainsi à la loyauté des traitements et à l'établissement de relations de confiance entre les organismes qui en sont responsables et les individus qu'ils concernent.

Ce principe s'applique à tout traitement de données personnelles, que les données soient :

- **directement recueillies auprès des personnes concernées :** par exemple, lors du renseignement en mairie d'un formulaire administratif, dans le cadre d'un entretien téléphonique, de l'ouverture en ligne d'un compte utilisateur, etc. ; ou
- **indirectement collectées :** collecte de données en libre accès sur Internet via le téléchargement de fichiers, le recours à des outils de moissonnage de données ou l'utilisation d'interfaces de programmation applicatives mises à disposition de réutilisateurs par les plateformes en ligne ; obtention d'informations auprès de partenaires institutionnels/commerciaux, réutilisation d'une base de données déjà constituée, etc.

Lorsque le responsable de traitement n'a pas directement collecté les données personnelles auprès des personnes concernées, il est dispensé de l'obligation d'informer ces personnes si cette information est impossible en pratique ou requerrait des efforts disproportionnés.

Quelles informations fournir et à quel moment ?

Lorsque l'information individuelle des personnes est matériellement possible, ce qui nécessite de disposer de leurs données de contact, le réutilisateur est en principe tenu de procéder à cette information. Elle doit avoir lieu aussi tôt que possible, et au plus tard lors de la première prise de contact avec les intéressés ou lors de la première communication des données à un autre destinataire s'il y en a une, et dans tous les cas **dans un délai ne dépassant pas un mois après la collecte.** Par ailleurs, lorsque l'information individuelle n'est pas possible, le réutilisateur doit en principe publier une information générale, collective (par exemple sur un site web), dans les meilleurs délais.

L'information doit comporter les points suivants :

- leurs **identité** et **coordonnées** (favoriser les différents modes de communication : adresses postale et électronique, téléphone, etc.), ainsi que les **moyens de contacter leur** délégué à la protection des données s'ils en ont désigné un ;
- la **finalité** et la **base légale de leur traitement** (voir la fiche n°2) **avec, le cas échéant, des précisions sur « l'intérêt légitime » le fondant** (p. ex. : l'intérêt de l'entreprise à mesurer son « e-réputation » à partir d'une analyse des commentaires la concernant publiés en ligne) ; doit également être précisée, si des données « sensibles » visées à l'article 9 du RGPD sont concernées, l'exception prévue par cet article (et, le cas échéant, du droit de l'Union ou de l'État membre en vertu duquel les données sont traitées) permettant de les diffuser (voir la fiche n°5) ;
- les **catégories de données utilisées** (p. ex. : identité, coordonnées, images, posts sur les réseaux sociaux ;
- les **destinataires ou catégories de destinataires** des données, avec, le cas échéant, des précisions quant au **transfert** envisagé de celles-ci vers un pays tiers à l'Union européenne.

En principe, il est également nécessaire de préciser :

- la **durée de traitement** des données, ou, lorsque ce n'est pas possible, les critères utilisés pour déterminer cette durée ;

- **la ou les source(s) accessible(s) au public** (p. ex. : sites web) **d'où elles proviennent**. **À noter :** lorsque plusieurs sources ont été exploitées et qu'il n'est pas possible d'attribuer les données d'une personne à une source en particulier, des informations générales sur ces sources doivent être fournies ; ces informations générales devraient toutefois préciser les différentes sources en cause lorsqu'elles sont connues (p. ex. : noms des différents réseaux sociaux sur lesquels les données auront été collectées) ;
- le cas échéant, **la mise en œuvre d'une décision entièrement automatisée,** les informations utiles à la compréhension de l'algorithme, **et les conséquences pour la personne concernée ;**
- l'existence de leurs **droits** (voir la fiche n°4), tel que le droit d'opposition (à mentionner clairement et séparément de toute autre information) si le traitement est fondé sur l' « intérêt légitime » ou la « mission d'intérêt public », ou de retirer son consentement à tout moment (si le traitement est fondé sur le consentement des personnes concernées).
- le droit d'introduire une réclamation auprès de la CNIL.

Comment la délivrer en pratique ?

Les personnes concernées ne doivent pas rencontrer de difficultés dans l'accès à l'information comme dans sa compréhension : elle doit être bien distinguée des autres indications sans lien avec la protection des données (support distinct des CGU), être aussi succincte et claire que possible (vocabulaire simple, phrases courtes, style direct, etc.) et adaptée aux conditions d'interaction avec les personnes.

Il existe ainsi **plusieurs moyens pour la fournir** : elle peut figurer sur le formulaire en ligne utilisé par le diffuseur des données pour recueillir celles-ci, ou transmise dans un document dédié en cas de collecte en face-à-face ; être mentionnée dans les courriels ou courriers adressés par un réutilisateur des données lors de son premier contact avec les personnes concernées ; être délivrée via un message vocal pré-enregistré, etc.

Pour atteindre l'objectif de concision et de bonne lisibilité, il **peut être procédé à une information en plusieurs niveaux**, **priorisant les informations essentielles que sont l'identité du responsable du traitement, les finalités et les droits des personnes.**

Exemple : ces informations sont données à la personne concernée directement sur la page d'inscription aux services proposées par une plateforme de communication en ligne, et sur cette même page, un lien renvoie vers une notice d'information complète.

À noter

En fonction du contexte, d'autres informations pourraient être délivrées prioritairement, telles que les caractéristiques importantes de leur réutilisation (p. ex. : réalisation d'un profilage, transmission des données à des partenaires commerciaux).

Quels sont les cas dans lesquels la délivrance d'une information individuelle n'est pas obligatoire ?

Dans certains cas prévus par les textes, les responsables de traitements peuvent ne pas procéder à une information individuelle des personnes concernées.

Pour déterminer si l'information individuelle des personnes est requise ou non, il convient de se référer aux fiches particulières déclinant les règles générales à certains types de réutilisation.

Cas n° 1 : la personne concernée a déjà obtenu les informations

Exemple

Cette hypothèse correspond au cas où les données ont été collectées sur Internet par un laboratoire de recherche pour un suivi de cohorte à l'année « n » ; les chercheurs les réutilisent pour les comparer avec les résultats de « n+1 » et les personnes concernées avaient bien été averties de la durée du suivi lors de la première collecte.

Attention

Le fait que, dans le cadre de leur CGU ou politique de confidentialité, des réseaux sociaux et autres plateformes de communication en ligne signalent la possibilité d'une réexploitation des données par des tiers ne permet pas à leurs réutilisateurs de considérer les personnes comme étant « déjà informées » : **l'information fournie doit en effet être particulière à chaque traitement des données mis en œuvre par chaque organisme concerné.**

Cas n°2 : la fourniture des informations rendrait impossible ou compromettrait gravement la réalisation des objectifs poursuivis

Par exemple, cette exception peut notamment concerner des travaux de recherche scientifique impliquant l'analyse de données diffusées sur les réseaux sociaux, si l'information des personnes concernées risque de les conduire à modifier leur comportement en ligne et, ainsi, de biaiser les travaux en question.

Attention

Lorsque cette exception apparaît mobilisable, **le responsable du traitement doit prendre des mesures appropriées** pour protéger les droits, libertés et intérêts légitimes des personnes concernées (ce point, qui concerne également le cas n°3, est précisé ci-dessous).

Cas n°3 : l'information se révèle impossible ou exigerait des efforts disproportionnés

Cet argument est souvent invoqué par les organismes privés qui réutilisent des données en libre accès sur Internet, dans la mesure où ils ne disposent généralement pas des données de contact des personnes dont ils collectent les données et où ceux qui les détiennent ne sont pas autorisés à les communiquer à toute personne souhaitant les exploiter. **Une analyse au cas par cas est à réaliser**, tenant compte du contexte spécifique de chaque traitement.

Concernant les efforts disproportionnés, deux cas de figure peuvent être schématiquement distingués.

Le réutilisateur des données dispose des coordonnées des personnes concernées, ou peut facilement et légalement y accéder

Dans cette hypothèse, le caractère disproportionné d'une information directe et individuelle est plus difficilement caractérisable.

Le caractère proportionné s'apprécie en mettant en rapport :

- **d'une part, l'atteinte portée à la vie privée des personnes dont les données sont traitées** ; par exemple, le fait d'utiliser des données publiques pour une étude interne à un organisme porte une atteinte moindre à la vie privée des personnes que le fait de les republier ou de les enrichir, en les rendant facilement accessibles ;
- **d'autre part, la difficulté et le coût d'une information individuelle**.
  - Lorsqu'une information par courriel est possible, elle est généralement requise.
  - Dans les autres hypothèses, une analyse au cas par cas est nécessaire, en prenant en compte notamment le coût de l'information, la faiblesse ou l'importance des risques que le traitement peut

faire encourir aux personnes, le fait que celles-ci peuvent ou non raisonnablement s'attendre au traitement de leurs données.

Il faut se garder de tout raisonnement automatique : ce n'est pas parce que les personnes sont très nombreuses que le responsable de traitement est systématiquement dispensé de l'obligation d'informer, il faut prendre en compte les autres paramètres et notamment l'intrusivité du traitement.

Exemples

- l'extraction par une entreprise des données de profils de réseaux sociaux (photos, établissements scolaires, employeurs, etc.) pour enrichir un service d'annuaire téléphonique ne peut légalement intervenir sans une information individuelle préalable (si 25 millions de personnes sont concernées, celle-ci ne constitue pas un « effort disproportionné » compte tenu du caractère intrusif du traitement et de la détention des coordonnées, a jugé le Conseil d'État – arrêt du 12 mars 2014, n° 353193) ;
- de même, et bien que les personnes puissent s'y attendre davantage, l'enregistrement par des cabinets de recrutement de données publiées sur les réseaux sociaux professionnels, à des fins d'enrichissement de bases de candidatures, doit s'accompagner de la délivrance d'une information à chacune des personnes concernées (voir la fiche n°14 du guide de la CNIL sur le recrutement).

À noter

Lorsque le coût s'attachant à la délivrance d'une information individuelle s'avère élevé et que l'exception des « efforts disproportionnés » ne semble pas pouvoir être mobilisée, le réutilisateur peut être amené à rechercher, voire à « investir » dans l'acquisition d'adresses mail.

Dans cette hypothèse, il doit impérativement veiller à ce que le recueil de ces données, sur internet ou auprès de courtiers en données notamment, ait pour seule finalité l'information des personnes concernées par son traitement et se fera dans le respect des règles de la protection des données (collecte entrant dans le champ des attentes raisonnables des intéressés, respect de leurs droits, suppression des données une fois l'information délivrée, etc.). De plus, la CNIL recommande que les données collectées se limitent à des coordonnées professionnelles.

À l'inverse, certains réutilisateurs pourront se prévaloir de l'exception des « efforts disproportionnés », notamment :

- si l'atteinte portée à la vie privée est particulièrement faible ;
- ou si les données ne sont pas conservées sous une forme permettant l'identification directe des personnes concernées.

Par exemple :

- j'ai préparé une note interne pour mon entreprise avec des éléments biographiques sur un parlementaire, à partir d'informations disponibles sur Internet : je n'ai pas besoin de l'informer de ce traitement, qui porte à sa vie privée une atteinte particulièrement faible, quand bien même je dispose de ses données de contact.
- une équipe de recherche réalise un projet de recherche scientifique sur les contenus du web social, intégrant des mesures de pseudonymisation ou d'anonymisation à bref délai ; ou une étude sur le fonctionnement des institutions démocratiques impliquant l'exploitation des listes publiques des personnes auditionnées par les parlementaires : ici encore, l'exception à l'information pourra être invoquée, compte tenu du nombre de personnes concernées et des contraintes importantes pour la recherche systématique de leurs coordonnées.
- une société spécialisée dans la représentation d'intérêts peut se prévaloir de la dérogation à l'information individuelle pour certains de ses traitements, ponctuels ou spécifiques, qu'elle met en œuvre à partir de données accessibles sur internet : ceux qui portent sur un sujet, une entreprise ou un secteur particulier et qui sont dits « pour comprendre », c'est-à-dire dont l'objectif n'est pas de contacter les personnes concernées (donc pas de recueil de leurs coordonnées) ; qui ne concernent que

des individus qui, de par leur activité, ont une forte visibilité dans l'espace public ; qui, enfin, présentent un faible degré d'intrusivité.

Le réutilisateur des données ne dispose pas des coordonnées des personnes, ou seulement d'anciennes informations, à l'exactitude incertaine (plus de 10 ans, par exemple).

Dans cette hypothèse, le caractère disproportionné d'une information individuelle pourra plus facilement être reconnu.

En particulier :

- lorsque les caractéristiques du traitement (finalité, portée, nature des données, garanties apportées, attentes raisonnables des personnes, etc.) ne l'impose pas (p. ex. : exploitation, à des fins d'études statistiques sur l'évolution des prix de l'immobilier, de la base de données « demandes de valeurs foncières / DVF » obligatoirement diffusée en *open data* par la direction générale des finances publiques et comprenant, à l'adresse, les transactions immobilières des cinq dernières années, sans identification directe des personnes concernées) ;
- lorsque l'atteinte portée à la vie privée par le traitement est particulièrement faible ;
- lorsque soit l'accès aux moyens de contact des personnes concernées ne paraît pas aisé ou pas souhaitable (p. ex. : personnes utilisant des pseudonymes, données pseudonymisées par le responsable du traitement initial), soit l'envoi induirait un coût trop lourd à supporter (p. ex. : nombre élevé de personnes et absence d'adresses mail facilement et légalement accessibles).

Attention

Afin de procéder à l'information, le réutilisateur sera parfois conduit à demander des coordonnées de contact à l'éditeur du site sur lequel il a récupéré les informations. Le fait que l'éditeur procède lui-même à l'information pour le compte du réutilisateur, qu'il y soit légalement tenu ou qu'il accepte d'y procéder, ne pose a priori pas de difficulté. En revanche, la transmission des coordonnées de contact au réutilisateur, qui constitue elle-même un traitement soumis au RGPD, n'est souvent pas légitime et nécessite le consentement des personnes.

À retenir

Tout réutilisateur de données souhaitant s'appuyer sur cette exception prévue à l'article 14.5 du RGPD doit :

- vérifier la réalité du caractère « impossible » ou « disproportionné » de la délivrance d'une information individuelle, en mettant en balance, dans ce deuxième cas, l'importance de l'atteinte à la vie privée (compte tenu des mesures envisagées pour la réduire : réduction du nombre de données collectées et de leur durée de conservation, garanties d'un niveau élevé de sécurité, réalisation d'une AIPD, anonymisation ou pseudonymisation des données, etc.), les efforts qu'une telle information requerrait et les effets que son absence pourrait avoir sur les personnes ;
- pouvoir en justifier à tout moment, notamment en documentant son analyse, conformément au principe de responsabilité.

Dans tous les cas, il doit également procéder systématiquement à la délivrance d'une information publique générale et complète, en n'hésitant pas à multiplier les supports de diffusion lorsque les personnes concernées sont nombreuses (publications sur le site web du réutilisateur, sur les comptes de réseaux sociaux qu'il utilise, dans un journal, etc.).

Cas n°4 : le traitement est mis en œuvre aux fins d'expression universitaire, artistique ou littéraire, ou d'exercice, à titre professionnel, de l'activité de journaliste

L'article 80 de la loi Informatique et Libertés prévoit que le droit à l'information peut être écarté pour les « *traitements mis en œuvre aux fins [...] d'expression universitaire, artistique ou littéraire* », lorsqu'une telle dérogation « *est nécessaire pour concilier le droit à la protection des données à caractère personnel et la liberté d'expression et d'information* ».

Une dérogation identique est prévue pour les journalistes professionnels (au sens de l'article L7111-3 du code du travail) qui se livrent à des activités d'investigation « en sources ouvertes ».

La protection des données personnelles doit en effet être conciliée avec la liberté d'expression et d'information. Une information minimale sur le traitement doit être fournie dans la publication, et en particulier l'identité du responsable de traitement. Cette exigence doit se combiner avec les règles spécifiques régissant les publications de presse.

Autres cas prévus par des dispositions légales constituant une mesure nécessaire et proportionnée pour garantir des objectifs d'intérêt public particulièrement importants

Par exemple

En application de l'article 48 de la loi Informatique et Libertés, le droit à l'information individuelle ne s'applique pas dans la mesure où une telle limitation est nécessaire au respect des fins poursuivies par ce traitement :

- aux traitements portant sur des données indirectement collectées et utilisées lors d'un traitement de l'État intéressant la sécurité publique ;
- aux traitements mis en œuvre par les administrations ayant pour mission de contrôler ou recouvrer des impositions, ou d'effectuer des contrôles de l'activité de personnes physiques ou morales pouvant donner lieu à la constatation d'une infraction ou d'un manquement, à des amendes administratives ou à des pénalités.

À noter

Les textes instaurant les traitements concernés et autorisant une telle dérogation doivent contenir des dispositions spécifiques, prévues par l'article 23 du RGPD, telles que les finalités du traitement en cause, l'identité de son responsable (ou l'identification des catégories de responsables), les catégories de données traitées et leur durée de conservation, les risques pour les droits et libertés des personnes concernées, ainsi que les garanties destinées à prévenir les abus.

Références

- Articles 12, 13 et 14 du RGPD
- Conformité RGPD : comment informer les personnes et assurer la transparence ?, cnil.fr
- L'information, design.cnil.fr
- Voir l'exemple de mentions d'information figurant dans la fiche « cas d'usage » dédiée à la réutilisation de données publiquement accessibles aux fins de diffusion d'annuaires de professionnels

Fiche n°4 : Quels sont les droits des personnes sur leurs données ?

Les personnes dont les données sont collectées en ligne à des fins de réutilisation, disposent de plusieurs droits sur celles qui les concernent. Ces droits vont leur permettre de conserver la maîtrise de leurs données et il appartient aux organismes responsables des traitements de les respecter et d'en faciliter l'exercice

De quels droits s'agit-il ?

En plus de devoir être informées de la mise en œuvre des traitements (voir à ce sujet la fiche n°3), les personnes concernées par les traitements de diffusion publique de données sur Internet et de réutilisation de données publiquement accessibles disposent de **différents droits qui constituent un ensemble de leviers d'action concrets leur permettant de contrôler les usages (objectifs poursuivis, conditions pratiques)** qui sont faits de leurs données.

Attention

Le choix de la base légale (voir à ce sujet la fiche n°2) **conditionne l'exercice de certains droits.**

Pour aider les organismes dans l'identification des droits applicables à leurs traitements, la CNIL a établi le tableau récapitulatif suivant. Les conditions d'exercice et de respect de ces droits, comme les cas dans lesquels il est exceptionnellement permis d'y déroger, sont détaillés plus bas.

Droits de la personne / Base légale	Accès	Rectification	Opposition	Effacement	Limitation	Portabilité
Consentement	✓	✓	✗ (1)	✓	✓	✓
Respect d'une obligation légale	✓	✓	✗	✗ (2)	✓	✗
Mission d'intérêt public	✓	✓	✓	✗ (2)	✓	✗
Intérêt légitime	✓	✓	✓	✓	✓	✗

(1) : à noter que si la personne concernée ne peut pas s'opposer, elle peut retirer son consentement au traitement.

(2) : à noter que les personnes concernées devraient toutefois pouvoir obtenir l'effacement de leurs données dans certains cas, en particulier si leur traitement est illicite ou n'est pas/plus strictement nécessaire.

En matière de réutilisation de données publiquement accessibles, les principaux droits applicables sont les suivants.

- **le droit d'accès** : les personnes peuvent demander à l'organisme la confirmation qu'il détient des données les concernant, en obtenir une copie pour en vérifier le contenu et solliciter certaines informations relatives aux caractéristiques du traitement mis en œuvre (finalités poursuivies, durées de conservation, identité des destinataires ou catégories de destinataires, sources des données, etc.) ;

> **Exemple**
> Une personne peut demander à un cabinet de recrutement s'il a collecté sur le web des données la concernant et, le cas échéant, solliciter la transmission d'une copie de celles-ci (dans un document qui les synthétise, notamment), ainsi que de toute information disponible sur les sites exploités et opérations réalisées (en particulier, existence éventuelle d'une décision entièrement automatisée et conséquences associées).

- **le droit de rectification : les personnes peuvent demander à l'organisme à ce que soient corrigées les données inexactes les concernant, ou à ce que soient complétées celles qui sont en lien avec la finalité du traitement ;**

> **Exemple**
> Un élu dont les données, diffusées sur des sites institutionnels, sont réutilisées par une association s'étant donné pour mission de faciliter l'accès au fonctionnement des institutions démocratiques, en particulier en établissant des statistiques sur l'exercice des mandats électifs, pourra demander à ce que soient corrigées les données statistiques les concernant et faisant l'objet d'une publication (p. ex. : taux de présence à l'assemblée, nombre d'amendements déposés, niveau de fidélité à la ligne de son parti).

- **le droit de retirer son consentement : lorsqu'un traitement de données est fondé sur le consentement préalable des personnes, celles-ci peuvent, à tout moment, sans justification particulière et via une modalité simple et équivalente à celle utilisée pour le recueillir, retirer leur consentement ;**

> **Exemple**
> L'utilisateur d'un site web de petites annonces relatives à des biens immobiliers, qui a mis en ligne une annonce en laissant ses coordonnées et qui a consenti à la réutilisation de ses données à des fins de prospection commerciale par des agences immobilières, doit pouvoir retirer son consentement quand il le souhaite (voir à ce sujet la fiche n°2).

- **le droit d'opposition : dans de nombreux cas (voir le tableau en page précédente), et notamment lorsque le traitement des données n'est pas fondé sur une obligation légale, les personnes peuvent s'opposer à tout moment, et pour des raisons tenant à leur situation particulière[6], au traitement de leurs données ; l'organisme doit alors cesser celui-ci, sauf s'il démontre qu'il existe des motifs légitimes et impérieux pour le poursuivre ou qu'il est nécessaire pour la constatation, l'exercice ou la défense de droits en justice ;**

[6] Le Conseil d'Etat a souligné dans un arrêt du 18 mars 2019 (n° 406313) qu'une personne se prévalant de son droit d'opposition ne pouvait « *se borner à invoquer des craintes d'ordre général concernant notamment la sécurité du fonctionnement de la base, sans faire état de considérations qui lui seraient propres* ».

Exemple

Une personne figurant dans les fichiers d'un organisme ayant pour activité la représentation d'intérêts (p. ex. : fichier de contact, cartographie des influenceurs), et dont les données ont été collectées en ligne (forums de discussion, réseaux sociaux, blogs, etc.) au titre de leur participation active au débat public sur des sujets cibles de l'organisme, doivent pouvoir s'opposer au traitement de leurs informations auprès de celui-ci (voir à ce sujet la fiche n°3).

À noter : si la finalité est la prospection, le droit d'opposition est discrétionnaire.

- **le droit à l'effacement : les personnes peuvent obtenir la suppression de leurs données, dans un certain nombre de cas, notamment quand ces données ne sont plus nécessaires au regard de l'objectif poursuivi, font l'objet d'un traitement illicite, sont traitées à des fins de prospection ou ont été collectées par une plateforme de communication en ligne, en particulier lorsqu'elles étaient mineures, ou encore quand ces personnes ont retiré leur consentement, ou se sont opposées au traitement et qu'il n'existe pas de motif légitime impérieux justifiant la poursuite celui-ci ;**

Exemple

Les courtiers en données (« *data brokers* ») spécialisés dans la collecte de données sur Internet, à des fins de commercialisation de celles-ci auprès d'entreprises ou de partis politiques, sont tenus de supprimer les informations se rapportant aux personnes qui en font la demande.

- **le droit à la limitation : les personnes peuvent demander à ce que leurs données soient temporairement « gelées » dans un certain nombre de cas, notamment lorsqu'elles exercent leurs droits d'opposition ou de rectification ; au cours du délai dont dispose l'organisme pour y répondre, celui-ci ne pourra pas les utiliser (sauf exceptions) ;**

Exemple

Une personne qui s'oppose à une certaine utilisation de données la concernant, recueillies sur internet par le responsable du traitement en cause, peut demander à celui-ci la cessation de l'exploitation de ses données le temps qu'il recherche d'éventuels motifs lui permettant de ne pas définitivement y mettre un terme.

- **le droit à la portabilité des données : lorsque les données sont traitées sur le fondement du consentement ou du contrat, les personnes peuvent recevoir, sous une forme directement exploitable, celles qui les concernent et qu'elles ont fournies au responsable de traitement, les réutiliser et/ou les transmettre à un autre responsable de traitement.**

Exemple

Une personne a consenti à la réutilisation, à des fins de prospection commerciale, de données la concernant publiées sur internet. Elle peut alors recevoir ces données dans un format pertinent (c'est-à-dire structuré, couramment utilisé et lisible par ordinateur) et même demander à ce qu'elles soient transmises directement à un autre organisme lorsque cela est techniquement possible.

Quelles sont les conditions d'exercice de ces droits ?

Ils doivent pouvoir être exercés sur simple demande, écrite ou orale, la personne concernée pouvant justifier de son identité par tout moyen.

Attention

Ne demander la fourniture d'une pièce d'identité que s'il existe un doute raisonnable.

Ce ne sera pas le cas, en particulier, si la personne exerce ses droits depuis un compte utilisateur créé au préalable, ou en utilisant la même adresse courriel que celle qu'elle a toujours utilisée pour ses contacts avec l'organisme.

De plus, l'exercice de l'ensemble des droits est gratuit. Toutefois, des frais raisonnables peuvent exceptionnellement être demandés dans certains cas d'exercice du droit d'accès (p. ex. : demande d'une copie supplémentaire).

Comment les respecter en pratique ?

L'organisme doit mettre en place un dispositif qui non seulement garantit l'effectivité des droits des personnes concernées, mais également facilite leur exercice.

Exemples de bonnes pratiques

- Fournir un formulaire de contact, un numéro de téléphone et/ou une adresse de messagerie dédié(e) à l'exercice des droits ;
- Si l'organisme dispose d'un site web intégrant des comptes utilisateurs, donner la possibilité aux personnes de les exercer à partir de leur espace personnel.

L'organisme doit garantir qu'il répondra aux demandes dans les meilleurs délais et, en principe, au plus tard dans un délai d'un mois. Ce délai peut être prolongé de deux mois en raison de la complexité de la demande ou du nombre de demandes que l'organisme aurait reçu de cette même personne. Dans ce cas, l'organisme doit informer la personne concernée des raisons de cette prolongation dans le délai d'un mois.

Recommandation : mettre en place une procédure interne prévoyant les conditions de gestion et de suivi des demandes d'exercice des droits.

Points de vigilance

Concernant le traitement des demandes de droit d'accès

Les données doivent être communiquées :

- **sous une forme électronique d'usage courant, lorsque la demande est faite en ligne, à moins que la personne concernée souhaite qu'il en aille autrement (p. ex. : fourniture orale ou par courrier postal) ;**
- **sous une forme compréhensible (les codes, sigles et abréviations utilisés doivent être explicités) ;**
- **via la transmission des supports originaux ou d'un document faisant la synthèse des données traitées (p. ex. : quelles informations pour quelles finalités) ;**
- **dans le respect des droits des tiers (secret des affaires, propriété intellectuelle, droit à la vie privée, secret des correspondances).**

Exemple : un cabinet de recrutement, qui collecte sur des réseaux sociaux professionnels les profils susceptibles d'intéresser certains de leurs clients, doit veiller à ce que les informations se rapportant aux personnes ayant pu commenter ces profils ne soient pas communiquées aux individus auxquels ces profils se rapportent.

Concernant le traitement des demandes de rectification, d'effacement et de limitation

À moins qu'une telle information soit impossible ou exige des efforts disproportionnés, l'organisme doit **informer chaque destinataire auquel il a communiqué les données des rectifications, effacements d'informations et limitations du traitement auquel il a procédé** en réponse aux demandes de personnes concernées (article 19 du RGPD).

En particulier, lorsqu'un diffuseur de données en ligne est tenu d'effacer les données qu'il a rendues publiques, il doit, en tenant compte des technologies disponibles et coûts de mise en œuvre, prendre les mesures raisonnables pour informer les tiers qui réutilisent ces données que la personne concernée a demandé l'effacement de tout lien vers ces informations, ou de toute copie ou reproduction de celles-ci (**article 17** du RGPD).

Une telle information peut être livrée par différents moyens, y compris techniques.

Pour en savoir plus, voir la fiche pratique suivante : https://www.cnil.fr/fr/webmaster-ou-responsables-de-sites-comment-repondre-aux-demandes-de-suppression-de-donnees.

Dans quels cas est-il possible d'y déroger ?

De façon générale, l'organisme peut refuser de donner suite aux demandes dont il est saisi lorsqu'il est en capacité de démontrer :

- **qu'elles sont manifestement infondées ou excessives, notamment en raison de leur caractère répétitif ;**
- **ou qu'elles se rapportent à des données ne permettant pas l'identification des personnes concernées.** Cependant, les personnes concernées pourront dans certains cas fournir des informations complémentaires permettant à l'organisme de leur « rattacher » des données et de faire ainsi suite à leur demande.

Attention

Si le responsable du traitement ne donne pas suite à la demande formulée par la personne concernée, il doit informer celle-ci, au cours du mois suivant la réception de la demande, des motifs de son inaction et de la possibilité d'introduire une réclamation auprès d'une autorité de contrôle et de former un recours juridictionnel.

De plus, certains droits connaissent des limites, qui vont découler de textes encadrant spécifiquement certains traitements (p. ex. : décret n° 2021-148 du 11 février 2021) ou directement du RGPD et de la loi Informatique et Libertés.

Ainsi par exemple, le droit à l'effacement ne s'applique pas dans certaines hypothèses, notamment quand le traitement des données est nécessaire :

- **à l'exercice du droit à la liberté d'expression et d'information**

Exemple

Dans certains cas, un journaliste peut refuser de supprimer des données figurant dans un de ses articles publiées en ligne au nom de la liberté dont il dispose en la matière, de même qu'un moteur de recherche peut refuser le déréférencement de certains contenus au nom du droit du public à accéder à l'information.

- **à la constatation, à l'exercice ou à la défense de droits en justice**

Exemple

Une entreprise qui, dans le cadre de la préparation ou de la gestion d'un contentieux, a collecté des données publiées en ligne pour disposer de preuves sur la diffamation publique dont elle aurait été l'objet, peut ne pas donner suite à une demande d'effacement.

- **à la réalisation de statistiques, ou de recherches scientifiques ou historiques,** dans les conditions prévues par le RGPD et sous réserve que l'application d'un tel droit rendre impossible ou compromette gravement la mise en œuvre de ces dernières.

Sous la même réserve que précédemment, **les finalités « statistiques » et « recherches » permettent également de déroger au droit d'opposition, au droit à la limitation, ainsi qu'au droit d'accès**, à condition, pour celui-ci, que les données soient conservées pour une durée limitée et sous une forme excluant manifestement tout risque d'atteinte à la vie privée (p. ex. : les personnes sont identifiées par un code attribué aléatoirement).

Point de vigilance

Dans le cadre de l'information individuelle ou générale devant être effectuée par l'organisme sur la mise en œuvre du traitement (voir la fiche dédiée), il convient d'indiquer aux personnes concernées le fait que leurs droits font l'objet de restrictions. Les motifs du refus de l'exercice d'un droit aux personnes concernées qui en ont fait la demande devraient également être expliqués, dans un langage compréhensible du grand public.

Références

- Articles 12, 15 à 21 du RGPD
- Articles 49 à 56 de la loi Informatique et Libertés

Fiche n°5 : Comment garantir la minimisation des données traitées ?

Le principe de minimisation : de quoi s'agit-il ?

Ce principe impose que les données personnelles traitées soient « *adéquates, pertinentes et limitées à ce qui est nécessaire au regard des finalités poursuivies* » (article 5.1.c du RGPD).

Pour garantir le respect de ce principe, en particulier au moyen des mesures développées ci-dessous, il **est essentiel que, préalablement à la mise en œuvre de leur traitement, les organismes qui réutilisent des données publiquement accessibles sur Internet :**

- **aient clairement défini l'objectif** poursuivi ;
- **se soient assurés de la légitimité de cette finalité** (art. 5.1.b du RGPD), notamment au regard de l'intérêt que le traitement présente pour des tiers ou pour eux-mêmes ; par exemple :
  - l'intérêt du public à connaître des résultats issus de travaux de recherche menés à partir de données librement accessibles sur internet ; ou
  - l'intérêt d'une entreprise à exploiter des données librement accessibles sur des réseaux sociaux professionnels à des fins de recrutement ;
- **identifient les catégories de données utiles et proportionnées** pour satisfaire cette finalité, et qu'ils pourront par conséquent légalement traiter ; ce, sous réserve dans certains cas du respect complémentaire de certaines dispositions spécifiques, comme celles relatives aux conditions de traitement des données sensibles (p. ex. : données se rapportant à la santé, à la vie sexuelle, aux opinions politiques, convictions religieuses, etc.) ;
- **définissent les mesures techniques et organisationnelles** propres à garantir la bonne application du principe de minimisation, et à pouvoir la démontrer conformément aux principes «*de responsabilité du responsable du traitement* » et de « *protection des données dès la conception et par défaut* » (art. 24 et 25 du RGPD).

En pratique : quelles mesures adopter pour le respecter ?

Lorsque cela est possible, ne collecter parmi les données en accès libre que celles qui sont indispensables à la réalisation de l'objectif poursuivi

Pour ce faire, les réutilisateurs de données doivent d'abord **mener une réflexion sur le champ des données « strictement nécessaires »**. Ainsi par exemple, dans le domaine « RH », les données utilisées par les recruteurs à des fins de constitution de profils de candidats doivent se limiter à permettre l'appréciation de leurs aptitudes professionnelles et capacité à occuper le poste proposé (voir à ce sujet la fiche n°14 du guide « Recrutement » de la CNIL).

Il leur faut également réfléchir aux modalités de leur collecte : lorsque celles-ci ne sont pas « manuelles », elles doivent être propres à garantir que le recueil d'informations n'excèdera pas en pratique le champ préalablement défini.

Lorsqu'il est impossible de calibrer précisément en amont le champ des données collectées (p. ex. : téléchargement d'une base de données en *open data* comportant plus de données que celles strictement nécessaires à la satisfaction du besoin du réutilisateur), **un tri doit être opéré le plus rapidement possible** entre les informations dont la conservation peut être justifiée et celles devant être supprimées.

- **Privilégier l'utilisation d'API[7] toutes les fois où de tels outils sont mis à disposition par les diffuseurs et de nature à répondre aux besoins**

[7] Voir la recommandation technique de la CNIL sur le partage de données par API (novembre 2023).

En plus de garantir, en principe, le fait que la collecte concernera des données se rapportant à des personnes ayant été mises en capacité de s'opposer à leur réutilisation (ou à certaines réutilisations) par des tiers, les API permettent de cibler davantage les données recueillies que le recours (par ailleurs souvent proscrit par les sites) à des outils de moissonnage des données (« *web scraping* ») mettant en œuvre des techniques d'extraction des contenus web, via des scripts ou programmes automatisés.

- **En cas de moissonnage de données**

  En plus de pouvoir compromettre fortement le bon fonctionnement des sites « moissonnés » (par exemple en occasionnant des dénis de service si trop de requêtes sont adressées simultanément), et être interdits par d'autres réglementations (par exemple par des CGU qui s'appuieraient sur le droit des producteurs de bases de données), les outils utilisés sont susceptibles, en fonction de la façon dont ils sont développés, de collecter toutes sortes de données. Il est donc nécessaire de veiller à :

  - **définir, en amont de la mise en œuvre du traitement, des critères précis et pertinents de collecte** afin d'opérer une distinction quant à la nature des données collectées ;
  - **limiter le risque de recueil de données sensibles**, lors de la sélection des mots clés ou catégories de contenus, par exemple ;
  - **supprimer toute donnée non pertinente, quel que soit son niveau de sensibilité, dès qu'elle est identifiée comme telle** dans le cadre de l'exploitation des données ainsi collectées (un tri exhaustif n'étant pas toujours possible de manière automatisée).

Anonymiser ou pseudonymiser les données quand leur exploitation sous une forme (directement) identifiante n'est pas requise

Toutes les fois où un traitement d'anonymisation ou de pseudonymisation des données collectées en ligne peut être mis en œuvre, il convient d'y recourir.

De tels traitements, qui excluent ou limitent le risque d'identification des personnes concernées, **doivent être envisagés aussi bien au stade de la collecte des données qu'au stade de leur exploitation**, notamment par les acteurs de la recherche scientifique. Ceux-ci peuvent utilement consulter la fiche pratique de la CNIL présentant les enjeux et avantages de ces techniques, ou encore sa délibération n° 2018-151 relative à un projet de traitement ayant pour finalité une recherche sur les impacts pour la vie privée des publications d'informations librement accessibles sur les réseaux sociaux.

À noter : à l'inverse et logiquement, lorsque de telles mesures protectrices ont été adoptées par le diffuseur des données, il convient de ne pas mettre en œuvre de traitements ayant pour objet ou effet de permettre la réidentification des personnes concernées.

Ne pas traiter de données « sensibles », sauf exception valablement mobilisable

Les réutilisateurs de données publiquement accessibles doivent se montrer très vigilants à l'égard des informations faisant l'objet d'une protection renforcée, telles que les données dites « sensibles ». Il s'agit en particulier de celles qui révèlent **la prétendue origine raciale ou ethnique, les convictions politiques, religieuses ou philosophiques ou l'appartenance syndicale, et des données concernant la santé, la vie ou l'orientation sexuelle.**

En effet, parce que leur collecte et exploitation soulèvent par nature des risques élevés pour les droits et libertés des personnes concernées, le RGPD (art. 9) et la loi Informatique et Libertés (art. 6 et 44) posent un **principe d'interdiction** de traitement de ces données. Ce principe s'accompagne toutefois d'exceptions, limitativement énumérées et d'interprétation stricte.

Ainsi, **si différents projets peuvent légitimement impliquer le traitement de données de cette nature** (p. ex. : mise en œuvre d'opérations ciblées de communication politique ou de lobbying, constitution d'une banque d'images/vidéos à des fins de recherche scientifique, réalisation d'une étude sur le comportement des personnalités politiques en matière de protection de l'intimité de leur vie privée), **que ce soit au stade de leur collecte en ligne ou par recoupements ultérieurs** (les données sensibles seront alors « inférées »), **leurs responsables doivent vérifier et être en capacité de démontrer, en particulier dans leur**

analyse d'impact relative à la protection des données (AIPD), qu'ils peuvent bien invoquer l'une ou l'autre des exceptions prévues par les textes, parmi lesquelles celles exposées ci-dessous.

- **L'exception « recueil du consentement »**

  Cinq critères cumulatifs doivent être remplis pour qu'un réutilisateur puisse s'en prévaloir : le consentement doit être libre, spécifique, éclairé, univoque et explicite. Il doit faire preuve de vigilance quant à ces conditions de validité et documenter les conditions de recueil de l'accord des personnes concernées (p. ex. : mise en place de mécanismes permettant d'obtenir une déclaration d'acceptation expresse et dédiée à cette question).

- **L'exception « données manifestement rendues publiques »**

  Dans un arrêt du 4 juillet 2023 (affaire C-252/21), la CJUE a souligné le fait que cette dérogation doit être interprétée de manière restrictive. Elle ne s'applique pas aux données concernant d'autres personnes que celles les ayant rendues publiques et il importe de vérifier si la personne concernée a entendu, de manière explicite et par un acte positif clair, rendre accessibles au grand public les données en question.

  Ainsi, lorsqu'un internaute insère des données sur des sites Internet ou applications en rapport avec les catégories de données sensibles visées à l'article 9 du RGPD, ou active des boutons de sélection intégrés à ces sites et à ces applications (ex. : boutons « j'aime » ou « partager ») ou les boutons lui permettant de s'identifier en utilisant les identifiants de connexion liés à son compte d'utilisateur du réseau social, son numéro de téléphone ou son adresse électronique, il ne peut être considéré comme ayant rendues manifestement publiques les données ainsi insérées ou résultant de l'activation de ces boutons que dans le cas où il a explicitement exprimé son choix au préalable, le cas échéant sur la base d'un paramétrage individuel effectué en toute connaissance de cause, de rendre les données le concernant publiquement accessibles à un nombre illimité de personnes.

  À noter : le fait que les données aient été manifestement rendues publiques par les personnes n'exonère pas les réutilisateurs de leur obligation de pouvoir justifier d'une base légale (voir la fiche dédiée à cette question).

- **L'exception de « nécessité pour la recherche publique »**

  Pour que des chercheurs puissent valablement se prévaloir de cette exception, ils doivent être en capacité de justifier, d'une part que la recherche en cause respecte certains critères précisés dans le code de la recherche (ce qui exclut la recherche privée) ; d'autre part, que ces utilisations sont rendues nécessaires « pour des motifs d'intérêt public important », après avis motivé et publié de la CNIL qu'il leur appartient donc de solliciter (pour plus de précisions sur le traitement des données sensibles dans le cadre de la recherche scientifique, voir la fiche cas d'usage dédiée à ce sujet).

À noter

Les réutilisateurs de données doivent également être vigilants à l'égard des informations qui, sans relever de l'article 9 du RGPD, présentent une sensibilité certaine, comme :

- **les données relatives à des infractions et condamnations**, qui ne peuvent être légalement traitées que par certaines catégories d'acteurs, et en particulier les réutilisateurs des décisions de justice pseudonymisées sous réserve que leurs traitements n'aient ni pour objet ni pour effet de permettre la réidentification des personnes concernées (art. 46 de la loi Informatique et Libertés) ;
- **ou encore celles relatives à des mineurs**, massivement présents en ligne et auxquels le cadre juridique de la protection des données accorde une attention particulière, s'agissant notamment des conditions de prise en compte de leurs droits.

Schéma récapitulatif : garantir la minimisation des données réutilisées

Ne collecter et/ou conserver que celles qui sont indispensables à la réalisation de l'objectif poursuivi

- ✓ **Mener une réflexion** sur le champ des données « strictement nécessaires »
- ✓ **Réfléchir aux modalités de leur collecte** : si celles-ci ne sont pas « manuelles », elles doivent être propres à garantir que le recueil d'informations n'excèdera pas en pratique le champ préalablement défini. Privilégier le recours aux API aux techniques de moissonnage de données
- ✓ **Opérer le plus rapidement possible un tri entre les informations dont la conservation peut être justifiée et celles devant être supprimées**, quand il est impossible de calibrer précisément en amont le champ des données collectées (p. ex. : téléchargement d'une base de données en open data comportant plus de données que celles strictement nécessaires à la satisfaction du besoin du réutilisateur)

Anonymiser ou pseudonymiser les données quand leur exploitation sous une forme (directement) identifiante n'est pas requise

- ✓ **Recourir si possible à un traitement d'anonymisation ou de pseudonymisation des données.** À envisager aussi bien au stade de la collecte des données qu'au stade de leur exploitation, notamment par les acteurs de la recherche scientifique
- ✓ **Ne pas mettre en œuvre de traitements ayant pour objet ou effet de permettre la réidentification des personnes concernées,** lorsque de telles mesures protectrices ont été adoptées par le diffuseur des données

Ne pas traiter de données « sensibles », sauf exception valablement mobilisable

- ✓ Principe d'interdiction de traitement des données qui **révèlent la prétendue origine raciale ou ethnique, les convictions politiques, religieuses ou philosophiques ou l'appartenance syndicale, et des données concernant la santé, la vie ou l'orientation sexuelle**
- ✓ **Quelques exceptions d'interprétation stricte**, parmi lesquelles :
  - Recueil du consentement
  - Données manifestement rendues publiques par les personnes concernées
  - Nécessité pour la recherche publique

Fiche n°6 : Comment garantir l'exactitude, la sécurité et la conservation limitée des données ?

Les principes à prendre en compte

Le RGPD impose que les données personnelles traitées soient « *exactes et, si nécessaire, tenues à jour* » (article 5.1.d).

Les diffuseurs et réutilisateurs de données doivent donc tenir compte de ce principe d'exactitude des données en veillant, tout au long de la vie de leur traitement, à ce que les informations qu'ils exploitent et qui sont erronées ou incomplètes soient effacées, rectifiées ou complétées sans tarder.

Le RGPD prévoit également que les données sont :

- conservées sous une forme identifiante pour la seule durée nécessaire à la satisfaction de l'objectif poursuivi par leur traitement ;
- **traitées dans des conditions de sécurité adaptées au niveau de risque** que le traitement présente pour les droits et libertés des personnes concernées.

Sauf à ce qu'il existe des dispositions légales s'y opposant (obligation d'archivage prévue par un texte), il convient ainsi de supprimer les données ou de les anonymiser dès qu'elles ont « rempli leur office » et, d'ici là, de les sécuriser autant que nécessaire pour garantir à la fois leur intégrité, confidentialité et disponibilité.

Conformément aux principes chapeaux de « *responsabilité du responsable du traitement* » et de « *protection des données dès la conception et par défaut* » (art. 24 et 25 du RGPD), les organismes diffuseurs et réutilisateurs de données « ouvertes » doivent prendre en compte les principes précédemment évoqués :

- **le plus en amont possible, c'est-à-dire dès les premières réflexions quant aux moyens de mise en œuvre de leurs projets de traitements ;**
- **au travers d'une identification des mesures techniques et organisationnelles à adopter pour garantir et démontrer en pratique leur bonne application.**

Ces mesures sont à déterminer au regard du contexte, de la portée et de la finalité de ces traitements, comme des risques qui leur sont associés pour les droits et libertés des personnes concernées.

La réalisation d'une analyse d'impact relative à la protection des données (AIPD), qui est obligatoire pour les traitements à risques élevés, a vocation à guider les responsables de traitements dans l'identification de ces mesures.

En pratique : quelles mesures adopter pour les respecter ?

Pour garantir l'exactitude des données traitées

Les réutilisateurs doivent prendre toutes les mesures raisonnables pour ne pas traiter des données obsolètes au regard des finalités poursuivies par leur traitement.

Il s'agit notamment, en cas d'exploitation de fichiers téléchargeables mis à disposition du public par les administrations sur « data.gouv.fr » ou d'autres espaces de diffusion en ligne, de tenir compte du rythme d'actualisation de ces fichiers tel que généralement mentionné, ou encore de mettre à jour régulièrement les informations « privées » collectées sur des plateformes en ligne, sachant que les enjeux pour les personnes concernées peuvent être très importants. Par exemple, l'absence de mise à jour des données collectées peut conduire à la rediffusion, par différentes sociétés réutilisatrices, d'informations que les personnes concernées avaient choisi de ne plus diffuser sur leurs profils publics de réseaux sociaux (voir à ce sujet les délibérations de la CNIL n° 2011-203 et n° 2012-156).

- **Recourir quand c'est possible à des API**, en ce qu'elles permettent de rafraîchir les données de manière régulière et automatisée, comme le souligne la recommandation de la CNIL n° 2023-050 dédiée à ces dispositifs.

- En cas d'utilisation de données de réseaux sociaux/blogs, en particulier à des fins de rediffusion de celles-ci, veiller au caractère très resserré des délais de mise à jour, compte tenu du fait que les préjudices d'image et de réputation sont susceptibles d'être d'autant plus grands que les informations et images en cause évoluent par nature très rapidement.

Pour garantir leur usage dans un temps limitée et des conditions sécurisées

Comme n'importe quel responsable de traitement, les réutilisateurs de données publiquement accessibles doivent veiller au respect des principes de sécurité et de durée limitée de conservation des données.

Notamment parce que la sécurité des données ne s'arrête pas à la préservation de leur confidentialité (l'article L322-1 du CRPA soumet ainsi la réutilisation des informations publiques *« à la condition que ces dernières ne soient pas altérées »* et que « *leur sens ne soit pas dénaturé* »), le caractère public des données n'a pas d'incidence sur les conditions de mise en œuvre de ces principes.

Dans ce contexte, les réutilisateurs de données publiquement accessibles sont invités à se reporter aux dossiers thématiques disponibles sur le site de la CNIL. Ils y trouveront de **nombreuses informations utiles quant à** la méthodologie à employer **et aux** mesures techniques et organisationnelles à adopter.

Fiche cas d'usage n°1 : la réutilisation de données publiquement accessibles aux fins de diffusion d'annuaires de professionnels

Cette fiche s'adresse aux organismes (associations, sociétés, etc.) qui publient en ligne, sous la forme d'**annuaires thématiques ou généralistes intégrant généralement des fiches / encarts / profils individuels**, des données professionnelles relatives à des personnes exerçant une activité commerciale, artisanale ou libérale.

Constitués de données personnelles au sens du RGPD[8], ces annuaires de professionnels **se distinguent des « annuaires universels »** (listes d'abonnés aux services de téléphonie fixe et mobile) régis par la directive « e-Privacy » et des dispositions spécifiques du code des postes et communications électroniques[9]. Ils sont **élaborés à partir de données publiquement accessibles et, notamment, de celles figurant dans les registres officiels** obligatoirement diffusés sur Internet dans un format ouvert.

Exemples de données publiquement accessibles

- Données de la base SIRENE des entreprises et de leurs établissements et du registre national des entreprises respectivement diffusés par l'Institut national de la statistique et des études économiques (INSEE) et l'Institut national de la propriété industrielle (INPI).
- Annuaires d'avocats publiés par le Conseil national des barreaux et les ordres des avocats des différents barreaux français[1].
- « Annuaire Santé » diffusé par l'Agence du numérique en santé et intégrant notamment le répertoire partagé des professionnels intervenant dans le système de santé (RPPS)[1].

Ces informations peuvent être **collectées par différents moyens** et, notamment, via des interfaces de programmation applicative (API) mises à disposition du public et/ou le lancement de requêtes automatisées sur les moteurs de recherche (avec indexation ou extraction de contenus). En plus de données d'identité et de contact professionnels, ces annuaires peuvent ainsi recenser des **données très diverses**, telles que la photographie des intéressés, leurs spécialités / domaines d'expertise, leurs diplômes, des références aux dossiers / sujets dont ils ont eu à connaître, à leurs travaux de recherche, des publications à leur nom ou les citant dans des magazines spécialisés, sur des réseaux sociaux, etc.

Par ailleurs, ces annuaires s'accompagnent généralement d'une **offre de services pour les professionnels référencés** (p. ex. : possibilité de « revendiquer » leur fiche et de souscrire un abonnement pour l'enrichir d'éléments les concernant et/ou d'être mis en relation avec des internautes via la plateforme)**. Les personnes ayant fait appel aux professionnels référencés disposent parfois de la faculté de publier des commentaires et de noter** la qualité de leurs prestations.

Si de tels annuaires sont susceptibles d'avoir des effets bénéfiques (visibilité pour les professionnels, information des consommateurs, etc.), ils comportent à l'évidence des **risques pour les droits et libertés des personnes référencées. L'exploitation de leurs données par des annuaires, combinées à un enrichissement potentiel de celles-ci (notamment par les notes et commentaires d'internautes), peuvent emporter une atteinte à leur vie privée ou à leur réputation**. **Ces risques se trouvent renforcés par le fait que ces données peuvent être inexactes et faire l'objet d'indexations** par différents moteurs de recherche externes.

Ces traitements de données font ainsi l'objet de **nombreuses plaintes auprès de la CNIL, en particulier de la part de médecins et d'avocats** qui, après saisie de leurs nom et prénom sur un moteur de recherche, découvrent l'existence de ces annuaires « parallèles ». Bien souvent en effet, les personnes concernées n'ont pas

[8] Les informations d'ordre professionnel restent des données personnelles, dont l'utilisation est soumise à la réglementation relative à la protection des données, comme souligné par la CJUE dans un arrêt du 9 mars 2017 (C-398/15). La CNIL l'avait auparavant rappelé dans sa délibération n° 2014-041 du 29 janvier 2014, confirmée par l'arrêt du Conseil d'État, 10ème SSJS, 30 décembre 2015, n° 376845 : « *Le nom et les coordonnées des personnes physiques, telles que leurs adresses et leurs numéros de téléphone, constituent des informations relatives à une personne physique identifiée et, par suite, des données à caractère personnel au sens des dispositions de la loi du 6 janvier 1978. Dès lors, que ces données soient des coordonnées professionnelles des personnes physiques en cause, et qu'elles soient le cas échéant par ailleurs rendues publiques, est sans incidence à cet égard* ». A noter toutefois que le RGPD ne s'applique pas aux noms, formes juridiques et coordonnées des sociétés, même lorsqu'elles incluent le nom d'une personne physique (ex : « Garage Michel Durand » SAS).

[9] Articles L34 et R10 et suivants du code des postes et communications électroniques (CPCE).

autorisé leur mise en œuvre, n'ont pas été informées de celle-ci et ne sont pas mises en mesure d'exercer leurs droits « informatique et libertés ».

Dans ce contexte, les développements qui suivent visent à **rappeler les principaux points de vigilance à considérer pour assurer la conformité au RGPD des traitements de données intervenant dans le cadre du référencement et, le cas échéant, de la notation des professionnels concernés.** Leur bonne prise en compte sera source de sécurité juridique et permettra le développement de relations de confiance, aussi bien avec les intéressés qu'avec les internautes exploitant le contenu de ces sites web.

Identifier la base légale des traitements

Lorsqu'ils n'interviennent pas exclusivement en vertu d'un contrat entre l'éditeur et les professionnels concernés (ce qui est généralement le cas dès lors que ces annuaires préexistent la plupart du temps à d'éventuelles relations contractuelles), **deux bases légales sont susceptibles de fonder ces traitements : l'« intérêt légitime » de la société éditrice (art. 6.1.f du RGPD) et le « consentement » des personnes concernées (art. 6.1.a du RGPD).**

De façon générale, la possibilité de recourir à la première est subordonnée au respect de certaines conditions : le traitement doit poursuivre un intérêt légitime de l'organisme ou d'un ou plusieurs tiers, être nécessaire à sa satisfaction et ne pas porter une atteinte excessive aux droits et intérêts des personnes concernées.

Pour savoir si la base légale de l'intérêt légitime peut valablement justifier un projet de collecte et d'exploitation de données à des fins de publication d'un annuaire en ligne, son responsable doit procéder à une mise en balance des intérêts en présence, **en appréciant tout particulièrement :**

- **les bénéfices attendus**, pour lui et les tiers, de la rediffusion de données publiquement accessibles et de leur éventuel enrichissement consécutif ;
- **les attentes raisonnables des personnes concernées à l'égard de tels traitements**, compte tenu notamment du contexte de la publication initiale de leurs données ;
- **les risques qui en résultent pour elles**, en particulier au regard de la nature des données en cause ; et
- **les garanties envisagées** pour écarter ou réduire les impacts potentiels.

Schématiquement, deux types de traitement peuvent être distingués :

1. Les annuaires constitués uniquement à partir de données diffusées en sources ouvertes (open data) à des fins de réutilisation : ces traitements peuvent se fonder sur la base légale de l'intérêt légitime.

Lorsqu'ils consistent uniquement à référencer des professionnels au sein d'un annuaire et se limitent, par défaut (c'est-à-dire sauf intervention directe de ces derniers), **à rediffuser les données « élémentaires » sur leur activité** (données d'identité, spécialités / domaines d'expertise, coordonnées du lieu d'exercice de la profession, etc.) qui se trouvent publiées dans un format ouvert en vertu d'un cadre légal spécifique (voir les exemples fournis en introduction), **leur licéité n'est pas subordonnée au recueil d'un consentement préalable du professionnel concerné**.

En effet, en favorisant utilement l'accès du public aux informations administratives et pratiques relatives à des prestations de services susceptibles de l'intéresser, ils s'inscrivent dans la continuité des objectifs poursuivis par les textes prévoyant la diffusion en sources ouvertes (*open data*) des registres / répertoires officiels de professionnels, au regard de la valeur qu'ils présentent pour la « communauté ».

Le fait qu'à la finalité informative de ces annuaires puisse s'ajouter une finalité commerciale (p. ex. : présence de bandeaux publicitaires à destination des internautes, dispositifs d'abonnements payants permettant aux professionnels de procéder à une présentation personnalisée de leur entreprise, d'être mis en relation avec les internautes via la plateforme, etc.), **est en principe sans incidence** sur la possibilité de recourir à la base légale de l'intérêt légitime : d'un côté, les sociétés éditrices sont autorisées à poursuivre un tel intérêt dans le cadre de leur activité à but lucratif ; de l'autre, les personnes concernées, en principe

préalablement informées de l'ouverture de leurs données par les administrations, pouvaient raisonnablement s'attendre au traitement en cause (principe de libre réutilisation des données publiques), et le risque d'atteinte à leurs droits et libertés est très limité compte tenu de la nature des informations traitées.

À noter

L'éditeur de l'annuaire devra prendre soin de mentionner la source des données qu'il diffuse et la date de leur dernière mise à jour (art. L. 322-1 du code des relations entre le public et l'administration - CRPA). Conformément au principe d'exactitude des données personnelles, il devra également veiller, autant que nécessaire, à leur actualisation (art. 5.1.d du RPGD), en utilisant le cas échéant l'API fournie par le diffuseur initial des données, ou en tenant compte de la fréquence de mise à jour du fichier à disposition du public en téléchargement.

2. Les annuaires qui viennent enrichir les données professionnelles en open data avec d'autres données publiquement accessibles ou fournissant un service de « notation » / commentaires : la possibilité de recourir à la base légale de l'intérêt légitime doit faire l'objet d'une appréciation au cas par cas.

De tels traitements sont par nature plus intrusifs et, suivant les conditions concrètes de leur mise en œuvre, présentent davantage de risques au regard de la protection des données.

L'analyse d'impact sur la protection des données, **qui devra souvent être menée préalablement à leur mise en œuvre** (AIPD, requise en cas de cumul de différents critères dégagés par le Comité européen de la protection des données : collecte à large échelle, surveillance systématique, croisement de données, profilage, évaluation / « *scoring* », etc.), **permettra notamment de déterminer s'ils sont propres à générer un niveau de risque pour les droits et intérêts des personnes qui ne saurait être considéré comme acceptable en l'absence de leur consentement préalable.**

S'agissant des traitements consistant à agréger des données venant de différentes sources

En vue d'assurer une plus-value à leurs annuaires, les éditeurs peuvent vouloir y intégrer d'autres informations que celles figurant dans les registres et répertoires officiels, en particulier lorsqu'il s'agit de les enrichir d'éléments en libre accès sur Internet (p. ex. : coordonnées électroniques, travaux de recherche, analyses publiquement exprimées sur un sujet, etc.).

Cet enrichissement sans le consentement préalable des personnes concernées est possible **si les collectes et traitements de données envisagés n'emportent pas une atteinte disproportionnée à leurs droits et intérêts, compte tenu notamment de leurs attentes raisonnables**. Pour apprécier l'existence d'un tel équilibre, doivent en particulier être considérés le contexte de la publication initiale des données (ex. : données diffusées sous une « licence ouverte » en vertu de dispositions légales), la nature de celles-ci et les conditions de leur collecte / enregistrement.

Par exemple, s'il s'agit de données dont la publication résulte d'une action positive des intéressés, pouvant être considérée comme témoignant, de façon manifeste, de leur souhait de voir ces informations associées à leur « identité numérique professionnelle » (p. ex. : leur numéro de téléphone mobile figurant à leur demande dans les annuaires publics d'abonnés à la téléphonie ; les informations concernant les expertise, expériences et compétences qu'ils mentionnent sur leur site web ou dans leur profil public d'un réseau social professionnel), leur rediffusion dans un annuaire généraliste ou thématique de professionnels présente moins le risque de heurter leurs intérêts que s'il s'agissait de données les concernant publiées par des tiers. Les éditeurs y procédant pourront ainsi se fonder plus aisément sur leur intérêt légitime, sous réserve de respecter les autres droits des personnes concernés (p. ex. : protection de certains contenus par le droit d'auteur ; pas d'exploitation de la photographie d'une personne sans recueil préalable de son consentement).

À noter

Il sera nécessaire de contrôler régulièrement, au moyen de mesures techniques et organisationnelles appropriées, que les données extraites des profils publics des réseaux sociaux professionnels n'ont pas fait l'objet de suppressions ou au contraire d'enrichissements par les personnes concernées.

D'autres traitements seront davantage de nature à créer un déséquilibre au détriment des droits et intérêts des professionnels référencés, en particulier ceux consistant à alimenter, au moyen de recoupements de données, la fiche de ces professionnels en éléments inférés ou dérivés (p. ex. : déduction de l'adresse électronique professionnelle à partir de celles de leurs collègues, établissement de corrélations entre elles et des personnes avec lesquelles elles ont collaboré). Ainsi, le consentement des personnes concernées devrait être recherché.

Enfin, certains traitements très intrusifs ne satisferont pas, à l'évidence, le test de la balance des intérêts. Ce sera par exemple le cas de l'enrichissement des fiches de l'annuaire en données sans lien direct avec l'activité professionnelle des personnes concernées, collectées notamment sur des plateformes de réseaux sociaux dédiées pour l'essentiel à des échanges d'ordre personnel. Une telle agrégation de données changerait la nature même de l'annuaire, en conduisant de fait à l'établissement de profils de *particuliers* au potentiel d'atteinte à la vie privée particulièrement élevé (voir à ce sujet la délibération de la formation restreinte de la CNIL n° SAN-2011-203 du 21 septembre 2011, confirmée par l'arrêt « Pages Jaunes Groupe » du Conseil d'État n° 353193).

À noter

Lorsque le consentement des personnes concernées est requis pour garantir la licéité d'un traitement, il devra, pour être valable, satisfaire aux conditions prévues par le RGPD, telles que rappelées dans cette fiche pratique.

Concernant les modalités envisageables pour son recueil, voir la fiche n°2 relative aux bases légales.

S'agissant des traitements intégrant la collecte et la publication de notes et commentaires d'internautes

La CNIL considère que, **sous réserve que leurs conditions de mise en œuvre ne portent pas une atteinte disproportionnée aux droits ou intérêts des artisans, commerçants et/ou professionnels libéraux référencés, de tels traitements pourraient se fonder sur la base légale de l'intérêt légitime** dès lors qu'ils participent à la liberté d'expression et d'information de l'éditeur et des contributeurs, concernent des professionnels vis-à-vis desquels le public dispose d'une « liberté de choix » (par opposition à d'autres catégories de professionnels, tels les enseignants de l'Éducation nationale)[10] et constituent une pratique expressément admise par le législateur (la loi n°2016-1321 pour une République numérique lui a donné un cadre juridique spécifique).

Il appartient toutefois à l'éditeur de l'annuaire de prendre toutes les mesures permettant de garantir l'équilibre de la balance des intérêts en cause et ainsi la licéité de son traitement.

À cet égard, la CNIL tient à souligner que **le fait d'assurer, aux professionnels qui le souhaitent, la liberté de ne pas/plus figurer dans l'annuaire commenté constitue, tout particulièrement, une mesure propre à garantir un tel équilibre**. En effet, il sera souvent difficile, pour les professionnels confrontés à des avis négatifs pouvant avoir de lourdes conséquences sur leur activité ou leur réputation, de faire valoir leur point de vue de façon circonstanciée et efficiente, en particulier du fait de la difficulté à identifier l'auteur du commentaire, de la lourdeur des procédures, parfois juridictionnelles, à mettre en œuvre pour obtenir la suppression de contenus abusifs, et de la subjectivité qui s'attache à l'appréciation de certains « savoir-être » et « savoir-faire ».

[10] En 2008, l'illégalité d'un site web (« Note2be ») organisant la notation de professeurs a été pointée par des juridictions (TGI Paris, ord. Réf. 3 mars 2008, n° 08/51650 et CA Paris, 14ème ch., sect. A, 25 juin 2008), ainsi que par la CNIL. Celle-ci a considéré que la société éditrice ne pouvait se prévaloir de son intérêt légitime compte tenu du fait, en particulier, que le site était susceptible de créer une confusion dans l'esprit du public avec un système officiel de notation. À l'inverse, dans un arrêt du 11 mai 2017, la Cour de cassation (pourvoi n° 16-13669, CNB c/ Jurisystem) a jugé que les tiers à la profession d'avocat ne sont pas tenus par les règles déontologiques de celle-ci, visant à garantir son indépendance, sa dignité et son intégrité, et qu'il leur appartient seulement, dans leurs activités propres, de délivrer au consommateur une information loyale, claire et transparente.

En l'absence d'une telle mesure d'opposition discrétionnaire, l'intérêt légitime requiert une analyse au cas par cas, en fonction notamment des incidences potentielles du traitement sur la situation des professionnels en cause.

A cet égard, l'éditeur pourra utilement se référer aux exigences posées par la norme *ISO 20488:2018* s'adressant aux gestionnaires d'avis en ligne de consommateurs (p. ex. : identification des auteurs d'avis et publication en premier des avis les plus récents).

En tout état de cause, il devra impérativement :

1. **Faire preuve d'une rigueur particulière pour informer, dans les conditions prévues par le RGPD et développées au point suivant, les personnes évaluées de la mise en œuvre du traitement, ainsi que pour garantir l'effectivité de leurs droits « informatique et libertés »**.

   À noter

   Bien que l'éditeur de l'annuaire puisse n'avoir, s'agissant des commentaires d'internautes, qu'un rôle passif (absence de contrôle avant publication) et ainsi le statut d'hébergeur de contenus au sens de l'article 6 précité de la LCEN, il doit être qualifié de responsable du traitement de ces commentaires au sens de l'article 4.7 du RGPD dès lors qu'il est à l'initiative de ce traitement et détermine les conditions de mise en œuvre de ce dernier (p. ex. : critères de notation et de classement des avis, durée de conservation et de publication). Par conséquent, il lui appartient, en particulier, d'examiner les demandes des personnes concernées exerçant leurs droits d'opposition et d'effacement.

   À noter que si le traitement est mis en œuvre à des fins de prospection, le droit d'opposition sera obligatoirement discrétionnaire (art. 21.3 du RGPD).

2. **Respecter les règles spécifiques du code de la consommation encadrant une telle pratique** (articles L. 111-7-2 et D. 111-17 et suivants), **en particulier la délivrance aux utilisateurs du site d'une information loyale, claire et transparente sur les modalités de publication, modération et de traitement des avis** (p. ex. : existence ou non d'une procédure de contrôle, critères de classement, date de « *l'expérience de consommation* » et durée de publication), **ainsi que la mise en place d'une fonctionnalité gratuite permettant aux responsables des produits/services de signaler un doute sur leur authenticité.**

À noter

Bien qu'elle ne soit pas exigée mais seulement encadrée par le législateur, la CNIL recommande la mise en œuvre d'une procédure de contrôle des avis intégrant, dans le respect des règles de protection des données personnelles, des mesures permettant d'identifier la source de l'information et de vérifier sa fiabilité : elle responsabilisera les utilisateurs du service et participera à la légitimité de l'éditeur de l'annuaire à se prévaloir du droit à l'information des internautes, comme au caractère adéquat, pertinent et non excessif des avis publiés.

Dans cette même perspective, la CNIL recommande également la mise à disposition :

- de critères de notation objectifs et appropriés au regard du type de prestation fournie par le professionnel (tenant compte par exemple du fait que les résultats obtenus par un médecin ou avocat ne dépendent pas nécessairement de la qualité de leurs prestations) ;
- des mentions de sensibilisation à l'intention des contributeurs, les alertant notamment sur la nécessité de ne pas faire un usage illicite ou abusif de leur liberté d'expression (p. ex. : injure, diffamation, dénigrement), ainsi que sur les risques encourus en pareil cas (sanctions civiles et pénales) ;
- d'un dispositif de signalement des propos répréhensibles, aux fins de suppression de ceux-ci par l'éditeur ou l'hébergeur.

Enfin, la CNIL recommande que la durée de conservation et de publication des avis n'excède pas cinq ans. Une durée raisonnable devrait être déterminée au cas par cas, en considération de la catégorie de professionnels concernés (ex. : les restaurateurs sont plus souvent notés que les notaires, et la qualité de leurs prestations est davantage susceptible d'évolutions dans le temps ; ces différences justifient qu'un délai plus court soit retenu pour les premiers que pour les seconds). À l'issue de ce délai, des données agrégées (par exemple, une agrégation des notes) pourraient être conservées et publiées selon les mêmes règles, avant d'être définitivement supprimées.

3. **Tenir compte des dispositions relatives au droit de réponse prévu par la loi n° 2004-575 du 21 juin 2004 pour la confiance dans l'économie numérique (LCEN – article 6.IV) et précisé par le décret n° 2007-1527.**

À noter

Si l'éditeur de l'annuaire dispose des coordonnées électroniques des professionnels référencés, une bonne pratique consiste en l'information systématique des personnes concernées par la publication d'un nouvel avis.

Attention

Les textes permettent le prononcé de sanctions aussi bien administratives que judiciaires, de nature civile comme pénale, pour non-respect des dispositions du RGPD, du code de la consommation et/ou de la LCEN, ainsi que pour les « délits de presse » prévus par la loi de 1881 (diffamation et injure, en particulier).

Informer les personnes concernées

Comme n'importe quel responsable de traitement, les éditeurs d'annuaires de professionnels sont concernés par le **principe de transparence des traitements de données personnelles** (voir la fiche n°3) : ils doivent informer les personnes dont ils traitent les données des objectifs qu'ils poursuivent et des conditions de mise en œuvre de leurs traitements. **Une telle information permet aux personnes de conserver la maîtrise des usages qui sont faits de leurs données**, en les mettant notamment en mesure d'exercer leurs droits.

Compte tenu des risques qu'une exploitation insuffisamment contrôlée des annuaires « enrichis » de professionnels fait peser sur les personnes (traitement de données non pertinentes, excessives ou inexactes ayant pour effet de porter atteinte à leur vie privée ou à leur réputation), les enjeux qui s'attachent au caractère effectif de leur information sont particulièrement importants.

Ainsi, ce n'est que dans des hypothèses très limitées que les éditeurs devraient pouvoir se prévaloir des dispositions de l'article 14.5.b du RGPD, autorisant les organismes ne collectant pas les données directement auprès des personnes concernées à ne pas informer celles-ci lorsqu'une telle information exigerait des « *efforts disproportionnés* ».

Deux situations peuvent être schématiquement distinguées.

Lorsque l'éditeur dispose des adresses électroniques des personnes, il devra, par principe, procéder à une information individuelle et complète de chaque professionnel concerné

En effet, dès lors qu'une information individuelle **peut être effectuée au moyen de l'envoi automatisé de courriels à chacune des adresses présentes dans la base de données, il est naturel de considérer que l'effort à fournir pour y procéder n'est pas « disproportionné » et ce, quel que soit le niveau de risque associé à la mise en œuvre du traitement**.

Une telle information devrait d'ailleurs poser d'autant moins de difficultés que l'éditeur cherchera souvent à contacter les professionnels pour les inviter, dans le cadre d'une action de prospection commerciale, à revendiquer leur fiche et à souscrire à des offres payantes (p. ex. : service de présentation personnalisée de leur entreprise, de mise en relation avec les internautes, etc.). En pareil cas, l'information fournie devra faire état de cette autre finalité ayant présidé à la collecte de leurs données.

Lorsque l'éditeur de l'annuaire ne dispose pas des adresses électroniques des professionnels, une analyse plus fine devrait être menée.

De façon générale, **l'éditeur pourra valablement se prévaloir de l'exception des « *efforts disproportionnés* » lorsque le traitement présente un caractère véritablement prévisible pour les personnes concernées**, et n'occasionne qu'un **niveau de risque faible pour leurs droits et intérêts** compte tenu de ses caractéristiques (ex : collecte et restitution de données brutes sans possibilité de publier un avis en ligne sur la qualité des prestations fournies).

Ce sera généralement le cas lorsque l'annuaire n'a vocation à exploiter que des informations publiques diffusées en *open data* en vertu du code des relations entre le public et l'administration ou de législations spéciales.

À titre illustratif, une association qui, dans le cadre d'actions visant à valoriser la composition du tissu artisanal de sa région, souhaiterait procéder à une « simple » rediffusion des données des professionnels concernés figurant dans le registre national des entreprises, serait légitime à arguer du caractère disproportionné de l'effort à fournir.

À noter

Dans ce cas, une information « RGPD » générale, complète et aisément accessible (distincte des conditions générales d'utilisation du site, en particulier), faisant notamment apparaître la source des données, devra tout de même être publiée sur le site web de l'éditeur. Dans l'idéal, cette réutilisation, parmi d'autres, ferait également l'objet d'une communication publique sur le site où sont diffusées en *open data* les données utilisées (sorte de « portail de transparence »).

En revanche, **s'agissant des annuaires enrichis d'autres types de données, et notamment des annuaires « commentés » de professionnels, plusieurs éléments laissent penser que les efforts à fournir pour les informer individuellement sont *a priori* proportionnés :**

- **de tels annuaires peuvent faire peser des risques sur les droits et intérêts des professionnels référencés** (vie privée, réputation), qui pourront être élevés compte tenu notamment de la volumétrie, de la nature des données traitées, de l'audience du site et des fonctionnalités proposées ;

- **en pratique, les éditeurs disposent toujours de la faculté de les contacter** (leur adresse administrative figure dans les registres officiels diffusés en sources ouvertes) ;
- **l'information constitue un élément clef de l'équilibre de la balance des différents intérêts en présence**, et le fait qu'elle soit fournie directement aux personnes concernées garantit qu'elles seront effectivement mises en mesure d'exercer leurs droits.

En toute hypothèse, l'information devra être délivrée aussi tôt que possible, au plus tard lors de la première prise de contact avec les intéressés ou de la communication des données à des tiers et, dans tous les cas, dans un délai ne dépassant pas un mois après la collecte. Elle devra l'être de façon « *concise, transparente, compréhensible et aisément accessible, en des termes clairs et simples* » (art. 12 du RGPD), et intégrer l'ensemble des éléments requis en cas de collecte indirecte des données (art. 14 du RGPD).

Pour plus d'informations à ce sujet, voir la fiche n°3 dédiée à l'information des personnes.

Exemple de mention d'information à l'intention des avocats référencés dans un annuaire « enrichi » publié sur Internet

Pourquoi le site « TousAvocats.fr » traite-t-il de mes données ?

Le site « TousAvocats.fr », placé sous la responsabilité de notre société X (située …), procède :

- **au référencement automatique de l'ensemble des avocats inscrits auprès d'un barreau français (création de fiches personnelles), sur la base d'une réutilisation d'informations publiques ;**
- **à la collecte et la publication d'avis relatifs à la qualité de leurs prestations, telle que ressentie par les personnes ayant fait appel à leurs services (notes sous forme d'étoiles assorties de commentaires).**

En tant qu'avocat en exercice, vous disposez d'une fiche dédiée, que nous vous invitons à consulter régulièrement, notamment pour vérifier l'exactitude des données y figurant. Vous pouvez y accéder suivant un critère matériel (spécialité), alphabétique (nom de famille) ou géographique (commune d'implantation), étant précisé qu'aucun classement « qualitatif » n'est effectué par notre société sur la base du calcul de notes moyennes.

Les traitements de données personnelles précédemment évoqués permettent, dans le cadre de la constitution et de la diffusion d'un annuaire « enrichi », la communication et l'obtention de nombreuses informations utiles et pratiques sur les modalités d'organisation et d'exercice par ses membres de la profession d'avocat. Ils se fondent juridiquement sur des « intérêts légitimes » au sens de l'article 6.1.f du règlement général sur la protection des données (RGPD), en ce qu'ils participent à la liberté d'expression et d'information de l'éditeur et des contributeurs, et sont réalisés dans des conditions respectueuses des intérêts, droits et libertés des personnes concernées.

Quelles sont les conditions de mise en œuvre de ces traitements ?

Votre fiche individuelle comporte des données relatives à votre identité, vos coordonnées professionnelles, votre ou vos spécialités et expériences professionnelles, ainsi qu'à tout élément relatif à la pratique de votre métier que vous avez souhaité publiquement valoriser (p. ex. : travaux de recherche, affaires traitées). Ces données proviennent :

- des sources publiques suivantes :

* base SIRENE de l'INSEE et annuaires officiels d'avocats publiés par le Conseil national des barreaux et les ordres des avocats des différents barreaux, diffusés en ***open data*** à des fins de réutilisation par des tiers ;

* sites web des personnes concernées et profils publics de réseaux sociaux professionnels (X et XX) ;

- des éléments que vous aurez souhaité apporter en complément, après revendication gratuite de votre fiche (procédure accessible ici : …).

Ces informations sont complétées de celles (notes sous forme d'étoiles et commentaires) qui résultent de l'exercice par les internautes de leur faculté de publier un avis en ligne, dans le respect notamment des

conditions légales encadrant une telle pratique, telles que rappelées par notre charte relative à la construction, la vérification et la modération des avis (charte accessible ici : ….).

Enfin, ces données, régulièrement actualisées (voir ici les conditions de leur mise à jour : …), sont conservées et publiées aussi longtemps que vous serez inscrit auprès d'un barreau français, à l'exception des avis en ligne dont le traitement n'excèdera pas cinq ans. Néanmoins, en vertu de vos droits d'opposition et d'effacement, les données vous concernant peuvent faire l'objet d'une suppression en amont.

Quels sont mes droits à l'égard de ces traitements et comment les exercer ?

Conformément au RGPD, vous disposez de droits d'accès, de rectification, d'opposition, d'effacement et de limitation (consulter ici la page dédiée aux droits : …), que vous pouvez exercer gratuitement et facilement à partir de votre compte en ligne, ou en contactant par courriel, courrier ou téléphone le délégué à la protection des données de notre société : […].

À noter que vous disposez également du droit d'introduire une réclamation auprès de la CNIL.

Respecter les droits des personnes

Les éditeurs d'annuaires devront enfin **garantir l'effectivité des droits « informatique et libertés » dont bénéficient les professionnels** concernés par leurs traitements (voir à ce sujet la sanction de la CNIL n° SAN-2021-014 du 15 septembre 2021). Ces droits constituent un ensemble de leviers d'action leur permettant de veiller à la préservation de leurs intérêts.

Le droit, pour les personnes référencées, de s'opposer à leur présence dans l'annuaire ou d'obtenir l'effacement de certaines informations les concernant

Tout professionnel référencé dans un annuaire en ligne privé, thématique ou généraliste, « enrichi » ou non et créé sur la base d'une réutilisation de données publiquement accessibles, dispose de la faculté d'exercer ses droits d'opposition et d'effacement.

Lorsqu'il s'oppose, « *pour des raisons tenant à sa situation particulière* », au traitement de ses données, l'éditeur de l'annuaire ne doit plus les traiter, « *à moins qu'il ne 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* » (art. 21.1 du RGPG).

Attention

S'agissant des annuaires comportant des notes et appréciations d'internautes, et comme indiqué précédemment, en présence d'une personne s'opposant de façon globale au référencement de ses données (et renonçant ainsi aux potentielles retombées positives que celui-ci pourrait avoir sur son activité), pour des raisons tenant à l'impact négatif fort de commentaires dont la véracité est contestée, la CNIL estime que l'éditeur du site devrait, généralement, accueillir favorablement la demande (interruption du traitement des données dans son intégralité). En effet, bien que le traitement en cause participe légitimement à la liberté d'expression et d'information de l'éditeur et des contributeurs, il semble pouvoir être estimé que les préjudices qu'il cause ou est susceptible de causer au professionnel doivent être prioritairement considérés par rapport au référencement d'un professionnel en particulier.

Exemple

Un médecin ou avocat dont la fiche comprend de mauvais commentaires, auxquels le secret professionnel lui interdit de répondre de façon circonstanciée, pourra obtenir la suppression de sa fiche au sein de l'annuaire au regard de son impact sur son image et sa réputation.

À noter

Si le traitement est mis en œuvre à des fins de prospection, le droit d'opposition sera obligatoirement discrétionnaire (art. 21.3 du RGPD).

De façon générale, les demandes de suppression portant sur certaines des données diffusées devraient faire l'objet d'une appréciation au cas par cas par le responsable du traitement, à la lumière des dispositions de l'article 17 du RGPD relatives au droit à l'effacement : analyse *in concreto* des motifs particuliers invoqués en cas d'opposition à leur traitement (pas de refus systématique et standardisé), du caractère illicite du traitement[11], etc.

Exemple

Toute personne dont les coordonnées postales professionnelles correspondent à sa résidence personnelle devrait pouvoir obtenir leur suppression de l'annuaire (le fait que cette information soit par ailleurs obligatoirement diffusée dans un registre officiel apparaît sans incidence, compte tenu de l'exposition supplémentaire que sa duplication emporte) ; de même, si la liberté d'expression et d'information peut mettre en échec une demande d'effacement, un professionnel pourra solliciter la suppression des commentaires d'internautes présentant un caractère diffamatoire, injurieux ou simplement excessif (sans lien avec leur pratique professionnelle ou ne satisfaisant pas, manifestement, l'objectif poursuivi par leur traitement, à savoir l'information du public).

Enfin, tout éditeur d'annuaires doit prendre des mesures raisonnables pour informer les principaux moteurs de recherche des demandes d'effacement auxquelles il donne une suite positive (art. 17.2 du RGPD).

Le droit d'accéder aux données les concernant et d'obtenir des informations sur les conditions de mise en œuvre de leur traitement

Tout professionnel référencé dans l'annuaire doit, en particulier, pouvoir obtenir, dans les conditions prévues à l'article 15 du RGPD, une copie des données le concernant.

Le droit de demander la rectification, la suppression ou l'enrichissement de toutes les données les concernant qui sont inexactes, obsolètes ou incomplètes

Exemple

Un professionnel en reconversion pourra exiger la suppression des données le concernant figurant dans un annuaire thématique relatif à son ancien secteur d'activité ; toute personne dont la carrière a évolué (nouvelles expériences professionnelles) pourra demander la prise en compte dans sa fiche des dernières évolutions.

Le droit de demander la limitation du traitement (cessation temporaire de la diffusion des données sur Internet)

Exemples

Le temps que l'éditeur de l'annuaire vérifie l'inexactitude de données mises en cause, ou, lorsque la personne s'est opposée au traitement de ses données, le temps qu'il analyse le point de savoir s'il dispose de motifs légitimes et impérieux prévalant sur ceux de la personne concernée.

---

[11] Dans un arrêt du 4 mai 2023 (affaire C 60/22), la Cour de justice de l'Union européenne (CJUE) souligne que le « traitement illicite » visé à l'article 17 du RGPD est le traitement mis en œuvre en « *violation par le responsable du traitement du principe de « responsabilité » tel qu'énoncé à l'article 5, paragraphe 2, dudit règlement, lu conjointement avec l'article 5, paragraphe 1, sous a), et l'article 6, paragraphe 1, premier alinéa, de ce dernier* » ; le traitement « *doit satisfaire aux conditions de licéité du traitement énumérées à l'article 6 dudit règlement. Par ailleurs, dans la mesure où les articles 7 à 11 du RGPD (…) ont pour objet de préciser la portée des obligations incombant au responsable du traitement en vertu de l'article 5, paragraphe 1, sous a), et de l'article 6, paragraphe 1, dudit règlement, le traitement de données à caractère personnel, afin d'être licite, doit également respecter, ainsi qu'il ressort de la jurisprudence de la Cour, ces autres dispositions dudit chapitre qui concernent, en substance, le consentement, le traitement de catégories particulières de données personnelles à caractère sensible et le traitement de données personnelles relatives aux condamnations pénales et aux infractions* ».

La bonne prise en compte de ces droits par les éditeurs d'annuaires suppose qu'ils aient mis en place :

- des **dispositifs visant à faciliter leur exercice** (p. ex. : fourniture de différents points de contact dédiés à la soumission des demandes ; faculté laissée aux personnes concernées d'agir directement sur la base de données, après vérification de leur légitimité à le faire) ;
- ainsi que **des procédures garantissant leur traitement effectif dans le délai maximal d'un mois** (art. 12 du RGPD).

Pour plus d'informations au sujet des droits des personnes concernées, voir la fiche n°4 dédiée à ce sujet.

Fiche cas d'usage n°2 : la réutilisation de données publiquement accessibles à des fins de constitution ou d'enrichissement de fichiers destinés à la prospection commerciale

La CNIL reçoit régulièrement des plaintes relatives aux pratiques de sociétés récupérant sur des sites web des données personnelles à des fins de prospection commerciale, alors même que certains des intéressés ont expressément indiqué s'opposer à une telle réutilisation de leurs informations.

Les sociétés visées ont généralement recours à des logiciels de moissonnage automatisé de données (ou *web scraping* en anglais) et sont notamment :

- des entreprises qui collectent pour leur propre compte et sur les annuaires en ligne l'ensemble des données personnelles se rapportant à leur secteur géographique, et qui démarchent ensuite les personnes concernées sur leur offre de biens ou de services ;
- des entreprises spécialisées dans la « pige immobilière », qui constituent des bases de prospects en moissonnant les coordonnées des personnes les ayant diffusées sur des sites de « petites annonces », dans le but de les louer / revendre auprès d'agences immobilières qui les utiliseront dans le cadre de leur recherche de nouveaux mandats de commercialisation[12].

D'autres sociétés exploitent des données publiquement accessibles pour enrichir leurs propres bases de prospects d'informations supplémentaires les concernant (p. ex. : ajout d'adresses courriel ou de données tirées des réseaux sociaux professionnels). Elles commercialisent également, à des fins de prospection ciblée, des profils de consommateurs sur la base de critères tels que la profession, la catégorie socio-professionnelle, la localisation, l'âge, le sexe ou encore les centres d'intérêt.

Bien que publiquement accessibles, les données en question sont des données personnelles. Dès lors, elles doivent être réutilisées dans le respect des dispositions du RGPD et, le cas échéant, des dispositions spéciales applicables, en particulier celles du code des postes et communications électroniques (CPCE).

Cette fiche pratique aborde les principaux points de vigilance et garanties que ces organismes doivent prendre en compte : les conditions de licéité des traitements, l'information des personnes concernées, le respect de leurs droits, l'encadrement de la sous-traitance et la nécessité potentielle d'une analyse d'impact.

Axée sur la réutilisation de données *publiées* sur Internet, elle ne traite pas de la question des *cookies* et autres traceurs à des fins de de ciblage publicitaire[13]. De plus, les conditions de mise en œuvre des actions de prospection commerciale faisant par ailleurs l'objet de nombreux contenus disponibles sur le site web de la CNIL, elle s'intéresse plus particulièrement à la phase de constitution et d'enrichissement des bases de données qui seront utilisées à de telles fins.

Garantir la licéité des traitements

Deux bases légales prévues par le RGPD sont susceptibles de fonder les opérations de collecte et d'exploitation de données en ligne à des fins de prospection commerciale (transmission des données à des partenaires commerciaux ou prospection par l'organisme ayant collecté les données) : le consentement des personnes concernées (art. 6.1.a) et l'intérêt légitime du responsable du traitement (art. 6.1.f).

Pour déterminer la base légale des traitements qu'ils mettent en œuvre, **les organismes doivent prendre en compte les éléments suivants.**

1) **Les traitements pourront se fonder sur la base légale du consentement lorsque les personnes auront donné préalablement leur accord à la collecte et à l'utilisation envisagée de leurs données, de manière libre, spécifique, éclairée et univoque.**

   Ce sera le cas, en particulier, lorsque :

[12] A noter que l'utilisation dans ce cadre de coordonnées est à distinguer de celle effectuée pour l'exécution de mandats, confiés par des particuliers, de recherche et de négociation de biens, laquelle ne poursuit pas une finalité de prospection commerciale au sens de l'article L34-5 du CPCE.

[13] Voir « Sites web, cookies et autres traceurs », publié sur cnil.fr

- **les personnes auront répondu positivement à une demande directe et précise de l'organisme, les ayant contactées via des coordonnées dont il disposait déjà licitement dans le cadre de la gestion de ses activités commerciales** ; le message suivant aura, par exemple, pu leur être adressé : « *Pour améliorer la connaissance de nos prospects et cibler davantage les offres commerciales qui leur sont adressées, nous souhaiterions collecter et exploiter les informations liées à leur situation géographique, profession et centres d'intérêts qu'ils ont choisi de diffuser, le cas échéant, sur leurs profils publics des réseaux sociaux suivants : …. Acceptez-vous que nous procédions à un tel enrichissement des données vous concernant* ? » ;
- **elles auront coché, sur le site où sont publiées leurs données, une case « J'accepte » (non pré-cochée) assortie de précisions sur les traitements en cause et l'identité des sociétés qui s'en prévalent** (nature des données exploitées, finalité de l'exploitation et noms des réutilisateurs autorisés) ; une telle situation suppose qu'un accord préalable ait été passé entre l'éditeur du site diffusant les données et les réutilisateurs de celles-ci.

Attention

L'acceptation de CGU informant l'utilisateur d'un site web de la possible utilisation de ses données à des fins commerciales, ou la validation sur le site d'une case à cocher par laquelle la personne concernée accepte, de façon générale et indifférenciée, la réutilisation de ses données à des fins de prospection commerciale, ne permettra pas à l'organisme de fonder son traitement sur la base légale du consentement.

En effet, la validité du consentement suppose notamment que l'identité des responsables de traitement qui s'en prévalent, pour une ou plusieurs finalités particulières, soit portée à la connaissance de l'utilisateur au moment de consentir.

Pour en savoir plus sur les conditions de recours à la base légale du consentement, voir la fiche pratique sur le site web de la CNIL.

2) L'utilisation licite de coordonnées publiées en ligne, à des fins de prospection par voie électronique et sans lien direct avec la profession de la personne concernée, est systématiquement soumise au recueil d'un consentement préalable.

Le consentement préalable des personnes concernées est requis lorsque l'exploitation envisagée par l'organisme consiste, au moins pour partie :

- **dans l'envoi par voie électronique (courriels, SMS/MMS, automates d'appels et fax) de messages publicitaires[14] ; ou**
- **dans la commercialisation, pour une utilisation à de telles fins, de coordonnées de particuliers**.

En effet, l'article L. 34-5 du CPCE interdit la mise en œuvre d'opérations de prospection par voie électronique en l'absence de consentement préalable des personnes concernées.

Si cet article ne s'applique pas aux traitements mis en œuvre par les courtiers en données (« *data brokers* » en anglais) lorsqu'ils ne prospectent pas eux-mêmes les personnes concernées, la CNIL estime que les transmissions de coordonnées auxquelles ils procèdent, et qui sont destinées à

[14] Ce point est souligné dans la décision de sanction pécuniaire prononcée par la CNIL à l'encontre d'une société exploitant les données de réseaux sociaux professionnels : n° SAN-2020-018 du 8 décembre 2020.

permettre de telles actions de prospection par leurs clients, doivent également se fonder sur le consentement préalable[15].

Pour autant, lorsqu'il apparaît possible de considérer, au regard de différents critères développés au point suivant (point 3), que les personnes concernées pouvaient raisonnablement s'attendre à l'utilisation de leurs coordonnées électroniques à des fins de prospection commerciale, la CNIL admet que les organismes, sur la base de leur intérêt légitime, procèdent à leur collecte en ligne en vue de recueillir le consentement requis de la part des intéressés.

Ainsi :

- une entreprise souhaitant prospecter des particuliers par voie électronique sur son offre de services, et ayant pu, à cette fin, collecter en ligne des adresses électroniques et numéros de téléphone mobile sur la base de son intérêt légitime, devra obligatoirement adresser un premier message aux personnes concernées dépourvu de tout caractère promotionnel et exclusivement destiné à recueillir leur accord à la réception de sollicitations commerciales par courriels et/ou SMS ; les données recueillies à cette fin devront être détruites en l'absence de réponse positive, à l'exception de celles permettant de tracer le refus des personnes concernées (soit leur réponse négative, soit leur absence de réponse) et éviter ainsi de les solliciter inutilement à l'avenir[16] ;
- de même, un courtier en données devra, avant de transmettre des adresses électroniques à ses clients aux fins de prospection par ceux-ci, recueillir l'accord préalable des personnes concernées en leur indiquant l'objectif de cette transmission et les catégories de destinataires ; il est recommandé de préciser, dans le message de demande, leur nombre approximatif et secteur d'activité ;
- enfin, les clients d'un courtier en données collectées en ligne devront également recueillir le consentement des personnes concernées avant toute opération de prospection par courriel ou SMS/MMS ; ce consentement pourra être recueilli soit par eux-mêmes dans le cadre de l'envoi d'un premier message « neutre » aux intéressés, soit par l'intermédiaire du courtier en données, dans le cadre de la fourniture aux personnes concernées, au stade du recueil de leur consentement, d'une liste précise et actualisée de ses clients.

Pour en savoir plus sur les modalités de recueil des consentements suivant les cas de figure, voir le référentiel de la CNIL « Gestion des activités commerciales »[17].

À noter

Sous réserve qu'elle entre bien dans le champ des attentes raisonnables des personnes concernées (voir le point 3 suivant), la collecte comme l'exploitation des coordonnées postales et téléphoniques à des fins de prospection par voie non électronique (envois de courriers postaux et appels téléphoniques hors automates d'appel), y compris la transmission de ces données à de telles fins, pourra se fonder sur la base légale de l'intérêt légitime du responsable du traitement.

Il en va de même des opérations de prospection par voie électronique visant exclusivement des personnes sollicitées au titre de leur activité professionnelle spécifique (prospection dite « entreprises vers entreprises »).

---

[15] Délibération n° 2021-131 du 23 septembre 2021 portant adoption d'un référentiel relatif aux traitements de données à caractère personnel mis en œuvre aux fins de gestion des activités commerciales, Légifrance.

[16] Voir la fiche pratique « Comment utiliser une liste repoussoir pour respecter l'opposition à la prospection commerciale ? ».

[17] Délibération précitée n° 2021-131.

Les droits des personnes concernées devront toutefois être impérativement respectés, comme précisé plus bas.

Pour approfondir :

- La prospection commerciale par courrier postal et appel téléphonique sur cnil.fr

3) **En l'absence de consentement préalable des personnes concernées, leurs données pourront être collectées sur la base légale de l'intérêt légitime uniquement s'il apparaît que cette collecte ne porte pas d'atteinte disproportionnée à leurs droits et intérêts.**

Selon le RGPD, le traitement ne peut être mis en œuvre sur la base légale de l'intérêt légitime lorsque les droits et intérêts des personnes concernées prévalent sur les intérêts de l'organisme. Celui-ci doit donc effectuer une pondération entre les différents intérêts en présence, tenant compte, tout particulièrement, des attentes raisonnables de ces dernières.

Cette prise en compte est essentielle : en l'absence d'un acte positif et explicite de la part des personnes concernées, l'intérêt légitime requiert de ne pas les surprendre dans les modalités de mise en œuvre comme dans les conséquences du traitement.

En amont de toute collecte de données en ligne, il appartient ainsi à l'organisme de considérer différents critères (méthode du faisceau d'indices) pour évaluer, au cas par cas, le caractère *a priori* acceptable et prévisible pour les individus de l'exploitation envisagée de leurs données.

Les critères à considérer portent notamment sur les points suivants :

- *La finalité poursuivie et la nature des données exploitées*

  Le caractère plus ou moins intrusif du traitement envisagé doit impérativement être pris en compte quand il s'agit d'analyser s'il n'heurte pas les droits et intérêts des personnes concernées et, en particulier, si celles-ci peuvent raisonnablement s'attendre à la collecte de leurs données dans l'objectif poursuivi par l'organisme.

  En effet, celles-ci s'attendront plus facilement à ce que soient collectées leurs coordonnées postales et/ou électroniques, éventuellement assorties de données d'identité et d'informations relatives à leur situation géographique et professionnelle, que d'autre catégories de données permettant d'établir des profils plus détaillés : par exemple, des informations relatives à leurs centres d'intérêts, préférences, ressentis, conditions / habitudes de vie et interactions sociales.

  À retenir

  Les traitements intrusifs visant à profiler des individus, suivant différents critères et sur la base de données collectées en ligne, aux fins d'établissement de profils et de ciblage publicitaire sont par nature porteurs de risques importants pour leurs droits et intérêts.

  En effet, les processus d'analyse des données publiquement accessibles sur les réseaux sociaux, blogs et forums, en vue de construire des profils permettant de mieux cerner la personnalité, les habitudes de vie, la situation sociale et/ou financière, les ressentis, relations, préférences et centres d'intérêts des personnes concernées, sont susceptibles d'aboutir à des conclusions et prédictions inexactes, voire à des décisions défavorables aux personnes, ainsi que de perpétuer des stéréotypes et d'enfermer les individus dans leurs choix (voir à ce sujet la fiche pratique sur le site web de la CNIL).

  Ainsi, leur licéité sera en principe subordonnée au recueil d'un consentement préalable compte tenu du caractère particulièrement intrusif du processus de profilage : agrégation de données concernant une même personne tirées de différents sites web, recoupements et inférences de données, etc.

Il faut également noter que la licéité d'opérations tendant à collecter ou à inférer des données sensibles (p. ex. : données révélant les convictions religieuses ou concernant la santé ou la vie sexuelle) requiert le recueil d'un consentement explicite, sauf s'il est possible de mobiliser une autre exception à l'interdiction de traiter des données sensibles (art. 9 du RGPD), ce qui apparaît peu probable. Pour plus de précisions à ce sujet (notamment concernant les éléments à prendre en compte), voir la fiche n°5 relative à la minimisation des données traitées.

- *La possibilité qui est ou non laissée aux personnes par le site source de s'opposer à la réutilisation de leurs données à des fins de prospection commerciale*

Le fait que l'éditeur du site web source ait prévu, sur le modèle de ce qui existe en matière d'annuaires universels (« liste orange »), un mécanisme d'opposition à ce type de réutilisations (mécanisme que l'on trouve par exemple sur certains sites de petites annonces de particuliers et professionnels) constitue un élément important à cet égard.

Il peut, par exemple, permettre à des sociétés spécialisées dans l'édition de piges immobilières de collecter les coordonnées des vendeurs de biens immobiliers ne s'étant pas expressément opposés à une telle exploitation de leurs données, compte tenu également du lien existant entre le contexte de la diffusion des données et l'objet des opérations de prospection envisagées.

Attention

La collecte des données se rapportant aux personnes ayant exprimé leur opposition à ce type de traitement, en cochant la case prévue à cet effet par le site éditeur ou en ne décochant pas une case qui permet une opposition par défaut, constitue un traitement déloyal de données personnelles interdit par le RGPD (article 5.1.a).

Il en va de même, en l'absence de mécanisme d'opposition mis à leur disposition par l'éditeur, lorsque les personnes ont manifesté d'elles-mêmes leur refus à être démarchées par des sociétés (p. ex. : insertion d'une mention dans leur petite annonce).

- *La nature du site web et les conditions d'accès aux données*

Le contexte de la collecte et la nature du site web qui diffuse les données doivent également être prise en compte.

Par exemple, il sera plus aisé de reconnaître le caractère prévisible de la collecte des données à des fins de prospection lorsqu'elles sont publiées sur un annuaire universel que sur un site de petites annonces spécifiquement dédié à la commercialisation, entre particuliers, de biens ou de services.

De même, devrait être considéré le fait que les données figurent ou non dans des contenus d'ordre essentiellement personnel d'ordre essentiellement personnel (ce n'est pas *a priori* pas le cas de ceux publiés sur un réseau social professionnel), ou encore que leur accès soit ou non subordonné à la création préalable d'un compte, devrait être considéré.

- *Le contenu de la politique de confidentialité ou des CGU du site éditeur*

La collecte des données publiées pour réaliser des opérations de démarchage est susceptible de ne pas entrer dans les attentes raisonnables des personnes lorsque la politique de confidentialité ou les CGU de l'éditeur :

  - interdisent le moissonnage automatisé ou la récupération manuelle des données à des fins de prospection commerciale ; ou
  - subordonnent la transmission de données au consentement préalable des personnes concernées.

- *L'objet de la prospection*

  La collecte des données à des fins de prospection commerciale peut davantage entrer dans les attentes raisonnables des personnes si elle concerne des biens ou services en lien direct avec :

  - l'activité professionnelle des personnes concernées (p. ex. : prospection par un organisme de formation en cybersécurité des délégués à la protection des données dont les coordonnées sont publiées en tant que données ouvertes (*open data*) par la CNIL ; prospection par une entreprise pharmaceutique des médecins référencés dans l'Annuaire Santé diffusé par l'ANS) ; ou
  - l'objet de leurs petites annonces (p. ex. : prospection des vendeurs de biens immobiliers par une agence immobilière, prospection des vendeurs de voitures par des garages).

  À l'inverse, si elle concerne des biens ou services sans aucun lien avec ces derniers (p. ex. : prospection des avocats, dont les coordonnées figurent dans l'annuaire en ligne du Conseil national des barreaux, par une entreprise de cosmétiques), il sera plus difficile de considérer que les personnes peuvent raisonnablement s'y attendre.

Lorsqu'il apparaît, au regard des critères exposés ci-dessus, que l'équilibre de la balance des intérêts n'est pas satisfait, notamment du fait que le traitement envisagé ne s'inscrit pas dans le cadre des attentes raisonnables des personnes concernées, l'organisme doit s'abstenir de procéder à la collecte des données publiées en ligne, y compris aux fins d'envoyer un premier courriel pour solliciter un consentement.

Dans le cas contraire, et comme vu précédemment, l'organisme pourra :

- **les collecter aux seules fins de recueillir le consentement des personnes concernées, lorsque leur exploitation le requiert** (cas de la prospection « entreprises vers consommateurs » par voie électronique obligatoirement fondée sur le consentement) ;
- **les collecter et les exploiter sur la base légale de son intérêt légitime**, lorsque la licéité de leur utilisation n'est pas assujettie au recueil d'un consentement préalable (cas de la prospection par voie non électronique ou effectuée en lien direct avec la profession des personnes concernées).

Dans tous les cas, ces opérations devront être effectuées dans le respect des autres dispositions du RGPD, en particulier de celles relatives à l'information et aux droits des personnes concernées.

Assurer la transparence des traitements et veiller à l'effectivité des droits des personnes concernées

Les exigences liées à l'information des personnes concernées

Tous les organismes doivent respecter le **principe de transparence des traitements de données personnelles** : même dans les hypothèses où la collecte des données en ligne puis leur exploitation pour de la prospection commerciale peuvent se fonder sur leur intérêt légitime, ils doivent informer les personnes concernées des objectifs qu'ils poursuivent et des conditions de leur mise en œuvre.

Cette information permet aux personnes concernées de **conserver la maîtrise des usages qui sont faits de leurs données**, en les mettant en capacité de consentir de façon éclairée au traitement envisagé ou, *a minima*, d'exercer leurs droits, notamment les droits d'opposition et d'effacement.

Cette information devra être délivrée dans les conditions prévues aux articles 12 et 14 du RGPD, à savoir :

- **en principe de manière individuelle** ;

- **aussi tôt que possible,** au plus tard lors de la première prise de contact avec les intéressés ou, le cas échéant, de la communication des données à des tiers, et, dans tous les cas, **dans un délai ne dépassant pas un mois** après la collecte ;
- **de façon concise, transparente, compréhensible et aisément accessible, en des termes clairs et simples**.

Son contenu devra **intégrer l'ensemble des éléments en principe requis en cas de collecte indirecte des données, et notamment** :

- les finalités et la base légale du traitement, comportant des précisions sur l'intérêt légitime mobilisé, le cas échéant ;
- les catégories de données collectées ;
- les destinataires ou catégories de destinataires de données ;
- la ou les sources d'où elles proviennent (lorsque plusieurs sources ont été exploitées et qu'il n'est pas possible d'attribuer les données d'une personne à une source en particulier, des informations générales sur ces sources doivent être fournies ; ces informations générales doivent toutefois préciser, lorsqu'ils sont connus, les noms des différents sites web sur lesquels les données auront été collectées) ;
- le cas échéant, l'existence d'un profilage, et des informations utiles concernant sa logique sous-jacente, ainsi que l'importance et les conséquences prévues pour la personne concernée ;
- les droits dont disposent les personnes concernées.

Les exigences liées aux droits des personnes concernées

Tout organisme collectant des données sur Internet aux fins de prospection commerciale, pour son propre compte ou par des tiers partenaires, doit :

- **tenir compte des oppositions « absolues » (indépendantes de tel ou tel traitement) exprimées par des personnes à une telle utilisation de leurs données** (personnes inscrites sur des listes anti-prospection, tel le dispositif BLOCTEL) ;
- **ne pas démarcher les personnes s'étant déjà opposées / ayant retiré leur consentement** à recevoir des sollicitations commerciales de sa part ;
- traiter les demandes d'exercice des **droits d'accès, de rectification, d'effacement et de limitation** (art. 15 à 18 du RGPD), dans les délais et conditions prévus par les textes ;
- **notifier aux partenaires commerciaux toute demande de rectification, effacement de données ou limitation du traitement** effectué (art. 19 du RGPD) lorsque les données leur ont été transmises ;
- **notifier à ces derniers toute demande de retrait du consentement effectuée, le cas échéant, auprès de lui** (article 7.3 du RGPD) lorsqu'il a recueilli le consentement pour les opérations de prospection réalisées par ses partenaires.

Pour plus d'informations :

- **Le principe de transparence des traitements de données personnelles**
- **Les droits des personnes concernées par les traitements**.

Réaliser si nécessaire une analyse d'impact sur la protection des données

Si le traitement projeté est susceptible d'engendrer des risques pour les droits et libertés des personnes concernées, au regard des critères identifiés par le Comité européen de la protection des données (p. ex. : traitement à large échelle, profilage, traitement de données sensibles, etc.), une analyse d'impact relative à la protection des données (AIPD) devra être obligatoirement réalisée avant la mise en œuvre du traitement (voir la rubrique dédiée du site de la CNIL).
Dans tous les cas, il s'agit d'une bonne pratique permettant à l'organisme de s'assurer, en liaison avec le délégué à la protection des données lorsqu'une telle personne a été désignée, que le traitement sera respectueux du RGPD.

Fiche cas d'usage n°3 : la réutilisation de données publiquement accessibles à des fins de recherche scientifique (hors santé)

Le web est source de quantités de données offrant des capacités inédites dans différents domaines de la recherche. Toutefois, le traitement de données personnelles publiquement accessibles, même pour une finalité légitime de recherche scientifique, reste soumis aux obligations du RGPD et de la loi Informatique et Libertés.

En complément **des fiches déjà publiées par la CNIL** sur la conformité des traitements de données à des fins de recherche (hors santé), cette fiche se concentre sur les spécificités des cas de réutilisation de données publiquement accessibles.

Elle n'aborde pas les traitements mis en œuvre à des fins de recherche, d'étude ou d'évaluation dans le domaine de la santé qui font l'objet d'un encadrement juridique spécifique, résultant des articles 66 et suivants de la loi Informatique et Libertés. Cette fiche reste toutefois applicable aux traitements de données de santé dans le cadre d'études ne relevant pas du domaine de la santé.

Qu'est-ce qu'une « recherche scientifique » au sens de la réglementation en matière de protection des données ?

La notion de « recherche scientifique » bénéficie d'une acception large dans le RGPD. Son considérant (159) énonce : « *Le traitement de données à caractère personnel à des fins de recherche scientifique devrait être interprété au sens large et couvrir, par exemple, le développement et la démonstration de technologies, la recherche fondamentale, la recherche appliquée et la recherche financée par le secteur privé* ».

Le Comité européen de la protection des données (CEPD), dans ses lignes directrices sur le consentement, donne un éclairage supplémentaire en indiquant que la recherche scientifique se caractérise par « *un projet de recherche établi conformément aux normes méthodologiques et éthiques du secteur en question, conformément aux bonnes pratiques* ».

La recherche a ainsi pour objet de produire des connaissances nouvelles dans tous les domaines dans lesquels la méthode scientifique est applicable.

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

- si le responsable de traitement remplit **un ou les deux critères organiques**, il peut bénéficier d'une présomption que son traitement relève de la **« recherche scientifique »** ;
- en revanche, s'il ne peut remplir **aucun des critères organiques**, il devra alors **examiner sa situation au regard de l'ensemble des critères de fond énumérés ci-dessous**.

Ces orientations d'interprétation ont pour vocation d'éclairer les responsables de traitement et ne présentent pas de caractère réglementaire.

Les critères organiques

Ces critères relatifs à la nature du responsable de traitement et celle du financement du projet constituent des indices, qui pourront être considérés comme suffisants en fonction des situations.

- **Critère n° 1 : la nature du responsable de traitement**

Exemple

Une université, un organisme de recherche tel que le CNRS, un centre de recherches comme l'Institut de recherche criminelle de la Gendarmerie nationale (IRCGN) ou une entreprise dédiant une grande partie de son activité à la « R&D » (dont les entreprises bénéficiant du statut de jeune entreprise innovante – JEI) ont vocation à mettre en œuvre ce type de traitement.

À noter

La qualification de responsable de traitement doit s'apprécier au cas par cas et en fonction de l'organisme s'il y en a un (v. les lignes directrices du CEPD sur les notions de responsable du traitement et de sous-traitant dans le RGPD).

En principe, en dehors du cas où le chercheur agit exclusivement pour son propre compte (chercheur « indépendant », c'est-à-dire sans rattachement à un organisme particulier et/ou sans mandat confié pour la réalisation de la recherche), le responsable du traitement mis en œuvre est l'organisme public ou privé au sein duquel il effectue sa recherche (université, entreprise privée, etc.). En effet, y compris lorsque le chercheur dispose d'une grande autonomie, l'organisme au sein duquel il travaille met à sa disposition des moyens de traitement pour réaliser des recherches, doit veiller à ce que les règles de protection des données personnelles soient respectées par ses chercheurs et peut pour cela donner des instructions aux chercheurs et en assurer le respect. Cela justifie que l'organisme assume la responsabilité des traitements mis en œuvre.

- **Critère n° 2 : le mode de financement**

Les projets de recherche peuvent être financés soit par des fonds publics, soit par des fonds privés, soit par des fonds mixtes.

Exemple

Dans le cas d'une entreprise privée responsable de traitement, le fait que le projet fasse l'objet d'une mesure d'incitation fiscale spécifique au soutien à la recherche scientifique (en particulier, le crédit impôt recherche – CIR) ou d'une aide publique à la recherche et au développement (par exemple, un financement de l'agence nationale de la recherche – ANR), l'indice pourrait être considéré comme suffisant.

Focus

La présomption de qualification de chercheur au sens de l'article 40-8 du règlement sur les services numériques

Un chercheur qui remplit les conditions fixées à l'article 40-8 du règlement sur les services numériques s'agissant de l'accès aux données des très grandes plateformes ou très grands moteurs de recherche aux chercheurs agréés, est présumé mettre en œuvre un traitement qui s'inscrit dans le cadre d'une finalité de recherche scientifique au sens du RGPD.

En effet, le règlement sur les services numériques exige plusieurs critères (parmi lesquels qu'ils soient affiliés à un organisme de recherche au sens de l'article 2-1 de la directive (UE) 2019/790, indépendants de tous intérêts commerciaux et que leur demande indique la source de financement des recherches, **etc.).**

Les autres critères de fond

À défaut de disposer d'indices suffisants au regard des critères nos 1 et 2 (notamment pour la recherche scientifique privée ne bénéficiant pas de financement public), il convient d'examiner conjointement les critères suivants (fondés sur le Manuel de Frascati et sur la définition de la R&D au sens du Crédit Impôt Recherche).

Ces critères étant cumulatifs, le responsable de traitement devra en principe démontrer qu'ils sont tous remplis pour que le traitement puisse être considéré comme relevant de la recherche scientifique au sens du RGPD. Lorsque ce n'est pas le cas, une analyse au cas par cas est nécessaire pour qualifier le traitement.

- **Critère n° 3.1 : la nouveauté** : Le traitement doit viser à obtenir des résultats nouveaux (une nouveauté pouvant aussi résulter d'un projet qui amène à constater des divergences potentielles avec le résultat censé être reproduit). L'objet de la recherche peut aider à la qualification de la recherche scientifique.

  À cet égard, la publication d'articles dans une revue à comité de lecture ou l'octroi d'un brevet permet de qualifier le critère de nouveauté.

- **Critère n° 3.2 : la créativité** : Ce critère repose sur des notions et hypothèses originales et non évidentes – l'apport des travaux à la connaissance scientifique ou à l'état de la technique. Le développement d'un savoir collectif qui ne profite pas seulement à l'entité morale porteuse du projet de recherche est un indice fort pour qualifier celle-ci de scientifique.

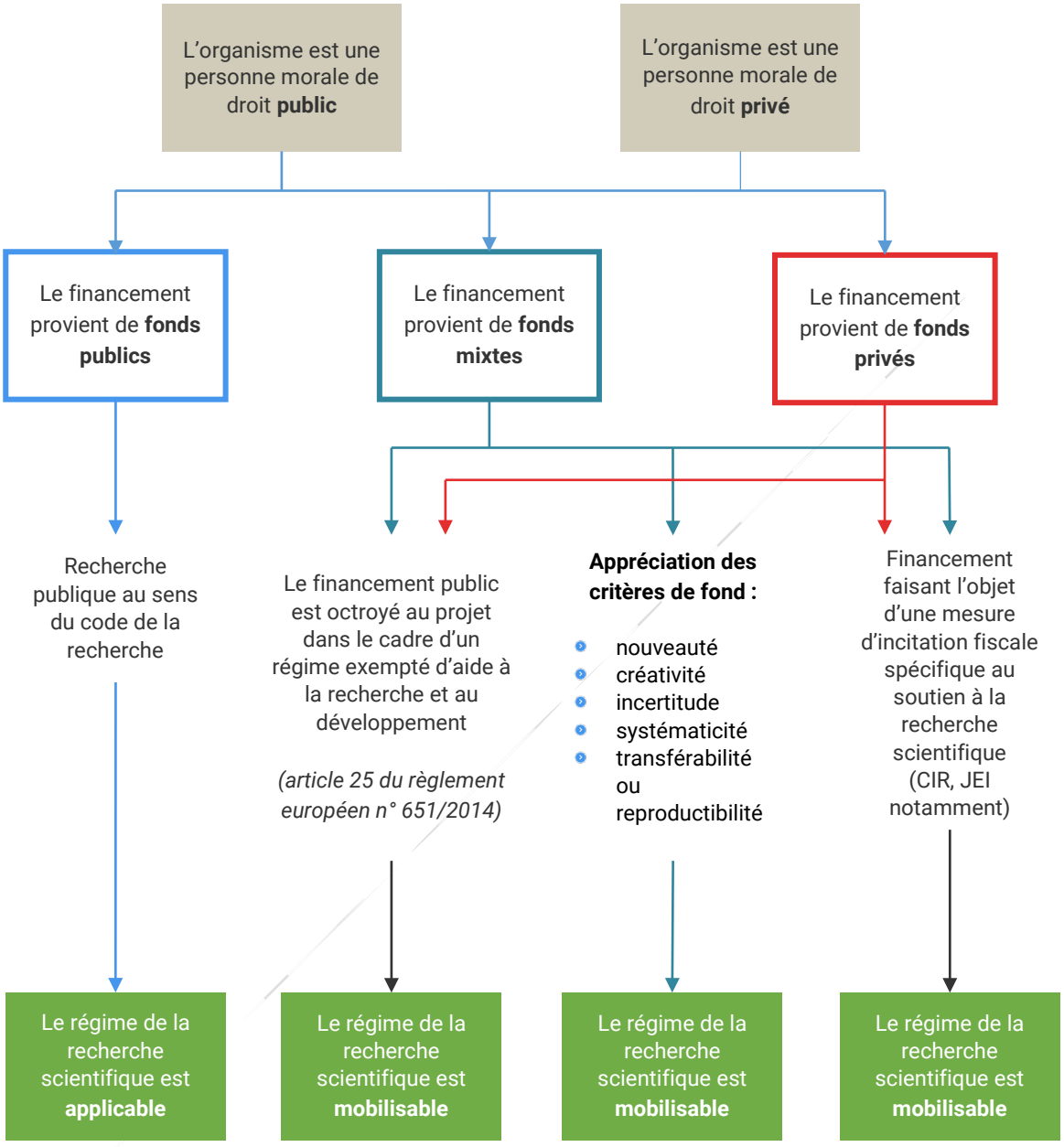
- **Critère n° 3.3 : l'incertitude** : Le traitement doit revêtir un caractère incertain quant au résultat final.

- **Critère n° 3.4 : la systématicité** : le traitement doit s'inscrire dans une planification et une budgétisation, et mettre en œuvre une méthodologie scientifique. Le respect de normes sectorielles pertinentes de méthodologie et d'éthique est un indice fort pour qualifier la recherche de scientifique. C'est par exemple le cas des exigences méthodologiques particulières pour les traitements mis en œuvre à des fins de recherche, d'étude ou d'évaluation dans le domaine de la santé qui résultent notamment des articles 72 et suivants la loi « informatique et libertés ».

- **Critère n° 3.5 : la transférabilité-reproductibilité** : le traitement doit déboucher sur des résultats qu'il est possible de reproduire ou de transférer dans un champ plus large que celui de la recherche.

  À titre d'exemple, la publication de l'étude réalisée et la présentation de la méthodologie de recherche adoptée est un indice fort permettant de souligner la volonté de partage du ou des porteurs de projet.

Schéma récapitulatif des critères de qualification de la « recherche scientifique » au sens de la réglementation en matière de protection des données



La licéité de la réutilisation des données

Le fait que le traitement poursuive un objectif de « recherche scientifique » au sens du RGPD ne suffit pas à garantir sa licéité.

Schématiquement, un réutilisateur de données personnelles doit d'une part se fonder sur une **base légale** au sens de l'article 6 du RGPD et, d'autre part sur une dérogation prévue par l'article 9 du RGPD s'il entend traiter des données sensibles.

L'exigence d'une base légale : généralités

En matière de recherche scientifique, les bases légales les plus pertinentes sont :

- **La mission d'intérêt public,** à condition que celle-ci soit prévue par une base juridique dans le droit auquel l'organisme est soumis (tel qu'une réglementation européenne, une loi ou un acte réglementaire), et que les traitements à des fins de recherche scientifique s'inscrivent dans cette mission.
- Si le responsable du traitement n'est pas investi d'une telle mission, **la base de l'intérêt légitime**, à condition qu'il prévale sur les intérêts, droits et libertés des personnes concernées. Sauf cas particulier, les personnes morales de droit public ne peuvent pas invoquer cette base légale.
- **Le consentement des personnes concernées au traitement de leurs données.** Bien qu'il soit difficile à recueillir s'agissant de la réutilisation de données publiquement accessibles, il sera souvent la meilleure base légale pour les traitements particulièrement intrusifs pour les personnes concernées.
- Lorsque cela sera possible (par exemple au moyen d'un outil de messagerie accessible par toute personne), le chercheur pourra solliciter directement auprès des personnes concernées le recueil de leur consentement pour traiter leurs données à des fins de recherche, en s'assurant que ce consentement est libre, spécifique et éclairé, qu'il peut être retiré à tout moment. Sinon, il devra solliciter le responsable du traitement de la publication initiale des données pour que ce dernier collecte un consentement valable en son nom (bien que cela suppose une coopération avec ce dernier qui n'est pas toujours possible en pratique).

L'exigence d'une base légale : cas d'application

Lorsqu'il s'agit de réutiliser des données publiquement accessibles à des fins de recherche scientifique, trois cas de figure sont à envisager :

- **Lorsque le chercheur entend réutiliser des données mises à disposition du public en vertu de la législation française ou européenne sur l'ouverture des données (*open data*)**, qui prévoit leur libre réutilisation.
- Dans ce cas, **le consentement ne sera presque jamais nécessaire**.

  Exemple

  Tel serait le cas si des chercheurs, salariés d'une ONG, menaient une recherche sur la liberté d'expression en France et envisageaient l'analyse d'un nombre important de décisions de justice mises à la disposition du public sur Légifrance (étant précisé que la recherche n'aurait ni pour objet ni pour effet de permettre la réidentification des personnes concernées, conformément aux articles 44.5 et 46.5 de la loi Informatique et Libertés). Ce traitement peut être fondé sur l'intérêt légitime poursuivi par l'ONG qui serait alors le responsable du traitement.

- **Lorsque les données n'ont pas été publiées dans une perspective d'*open data*** mais pour une finalité particulière (par exemple les données figurant sur les réseaux sociaux ou des forums de discussions spécialisés). Dans ce cas, **une analyse au cas par cas sera requise pour s'assurer que le traitement ne porte pas une atteinte disproportionnée aux intérêts, droits et libertés des personnes concernées**.

  En particulier, il s'agira de prendre en compte :

  - le caractère plus ou moins privé des informations recueillies ;

- les catégories de personnes (p. ex. : personnalités publiques, mineurs) et de données concernées, plus ou moins porteuses de risques (p. ex. : données professionnelles ou données relatives à la vie privée),
- les mesures susceptibles d'être adoptées pour limiter l'impact du traitement sur les personnes décrites ci-dessous (p. ex. : anonymisation à bref délai, pseudonymisation, droit d'opposition inconditionnel).

Exemples

- Des chercheurs au sein d'un laboratoire public étudient l'évolution de la langue française sur Internet. Ils décident pour cela de collecter et d'analyser des commentaires publiés librement sur différents réseaux sociaux qu'ils anonymisent à bref délai pour suivre le nombre d'occurrences de certaines expressions ou formes orthographiques. Dans la mesure où le responsable de traitement est un laboratoire public, la base légale de la mission d'intérêt public pourrait être mobilisée.
- Un institut de recherche privé mène une recherche scientifique pour évaluer l'impact sur l'emploi de l'IA générative pour la génération de code informatique, et envisage de collecter et d'analyser des données publiées sur un réseau social professionnel (par exemple, des profils publics, des commentaires et publications librement accessibles). L'intérêt légitime peut être envisagé selon les cas comme une base légale valable à condition que des mesures suffisantes soient prises (telles que l'anonymisation des données lorsque cela est possible, leur pseudonymisation à défaut, ou encore telles que des mesures de sécurité des données).

- **Enfin, il arrive que des jeux de données contenant des données personnelles soient librement mis à disposition sur Internet en dehors de toute législation**. Le plus souvent, il s'agit de données qui étaient déjà publiquement accessibles et qui constituent une base de données ou un corpus diffusé sur le site web d'une université ou d'une plateforme dédiée au partage de jeux de données, pour faciliter leur réutilisation.

  Dans ce cas spécifiquement, et de façon plus générale, le contrôle de la licéité de la mise en ligne de la base de données relève en premier lieu du responsable de traitement qui opère cette mise en ligne. Cependant, **l'opérateur qui réutilise les données doit s'assurer qu'il n'est pas en train de réutiliser un jeu de donnés dont la constitution était manifestement illicite** (par exemple provenant d'une fuite de données).

  En effet, la CNIL estime qu'il résulte du principe général de licéité des traitements de l'article 5.1.a du RGPD que le réutilisateur ne saurait réutiliser un jeu de données dont il ne peut ignorer qu'il est constitué ou mis en ligne en méconnaissance du RGPD ou de règles plus générales (telles que celles interdisant les atteintes à la sécurité des systèmes d'information ou des atteintes à des droits de propriété intellectuelle).

  En outre, la personne qui télécharge ou réutilise un jeu de données manifestement illégal risque de se rendre coupable du délit de recel (article 321-1 du code pénal).

  Si la possibilité de réutiliser un jeu de données librement mis à disposition sur Internet n'est pas subordonnée à des vérifications approfondies sur le respect de l'ensemble des règles du RGPD ou d'autres règles juridiques applicables (droit d'auteur, données couvertes par le secret des affaires, etc.), vérifications qui relèvent en premier lieu de l'opérateur qui met en ligne les données, **la CNIL recommande aux réutilisateurs de s'assurer :**

  - **que la description du jeu de données mentionne leur source ;**

Exemple

Sur un blog librement accessible sur Internet, un chercheur identifie un jeu de données comportant un ensemble de discussions entre particuliers. Selon sa description, les données sont pseudonymisées mais aucune source n'est mentionnée. Le chercheur devrait alors s'abstenir de le réutiliser sans obtenir davantage de précision qui lui permettrait de lever ses doutes quant à la conformité de sa constitution et de sa diffusion.

- que la constitution ou la diffusion du jeu de données **ne résulte pas manifestement d'un crime ou d'un délit ou a fait l'objet d'une condamnation ou d'une sanction publique** de la part d'une autorité compétente qui a impliqué une suppression ou l'interdiction d'exploiter ultérieurement les données ;
- que la constitution ou la diffusion de la base de données ne résulte pas manifestement d'un crime ou d'un délit ou a fait l'objet d'une condamnation ou d'une sanction publique de la part d'une autorité compétente qui a impliqué une suppression ou l'interdiction d'exploiter ultérieurement les données ;

Exemples

- En cas d'achat d'un jeu de données sur le *dark web* provenant par exemple d'une atteinte à un système de traitement automatisé punie par la loi (au sens de l'article 323-1 du code pénal), un chercheur ne saurait en ignorer l'origine délictuelle et dont l'illicéité serait alors manifeste.
- Il en irait de même pour un chercheur souhaitant réutiliser un jeu de données pour lequel une décision de justice a retenu une atteinte à un droit de propriété intellectuelle comme le droit *sui generis* des producteurs de bases de données (au sens de l'article L. 343 du code de la propriété intellectuelle).

- que l'origine des données soit suffisamment documentée pour qu'il n'y ait **pas de doutes flagrants sur la licéité** du jeu de données (notamment que le traitement source ne soit pas manifestement dépourvu de base légale lorsque les données sont tellement intrusives qu'elles ne sauraient être traitées sans le consentement des personnes) ;

Exemple

- Sur une plateforme d'échange de jeux de données, un chercheur repère un ensemble compilant les trajets domicile-travail de milliers de personnes. Sa description explique qu'il s'agit de données de géolocalisation précises, non anonymes. Dans cette hypothèse, le chercheur ne saurait ignorer qu'il existe un doute sérieux quant à la licéité de la diffusion d'un tel jeu de données en l'absence de consentement des personnes.
- À l'inverse, un chercheur pourrait réutiliser un jeu de données dont la description ne laisse pas de doute flagrant quant à sa licéité. Il en irait ainsi d'un jeu de données pseudonymisées, initialement rendues publiques par les personnes concernées sur un site web identifié et qui ne contiendrait pas de données sensibles.
- Il en irait de même pour la réutilisation d'un jeu de données dont le diffuseur les présenterait comme anonymes, par exemple s'il s'agit de données agrégées provenant des utilisateurs de son service.

- que le jeu de données ne contienne **pas de données sensibles ou de données d'infraction** (au sens des articles 9 et 10 du RGPD), ou de mener des vérifications supplémentaires pour s'assurer que ce traitement était licite (il s'agirait principalement de s'assurer du recueil d'un consentement explicite des personnes concernées, ou que les données ont été manifestement rendu publiques par ces dernières comme cela est précisé ci-dessous).

Exemple

Sur un forum en ligne, un chercheur découvre l'accessibilité d'un jeu de données non anonymes qui contiendrait, selon sa description, les parcours de soin d'une centaine de patients atteints d'une pathologie particulière qui proviendrait d'hôpitaux français. Dans ce cas, le chercheur devrait sérieusement douter de la licéité de la diffusion de ce jeu de données compte tenu de l'encadrement des données sensibles prévu par le RGPD et la loi Informatique et Libertés.

Ces vérifications préalables pourraient utilement figurer dans l'analyse d'impact relative à la protection des données (AIPD).

Il est à noter que certains manquements commis par le responsable des traitements de constitution et de diffusion d'un jeu de données n'impactent pas directement et irrémédiablement la licéité des

traitements mis en œuvre par le réutilisateur, à condition que la réutilisation soit elle-même conforme au RGPD.

Exemple

Tel serait notamment le cas de la fourniture de mentions d'informations incomplètes lors de la constitution ou de la diffusion du jeu de données[1], ou du défaut de documentation adaptée de la conformité de ces traitements (par ailleurs souvent difficile à vérifier sans interaction avec le diffuseur ou l'éditeur du jeu de données).

En tout état de cause, ces vérifications préalables n'exonèrent pas les réutilisateurs d'une analyse complète de la conformité de leurs propres traitements, y compris lorsqu'ils réutilisent des jeux de données dont la constitution et la diffusion ne relèvent pas du droit français ou européen (contrairement à leur réutilisation par une entité établie sur le territoire français ou européen qui est soumise au RGPD).

Les dérogations permettant de traiter des données sensibles

Pour traiter des données sensibles, outre une base légale valide, des conditions supplémentaires doivent être remplies.

Les données sensibles (au sens de l'article 9 du RGPD) sont celles relatives à l'origine ethnique ou prétendument raciale, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale, ainsi que le traitement des données génétiques, des données biométriques[18] aux fins d'identifier une personne physique de manière unique, des données concernant la santé ou des données concernant la vie sexuelle ou l'orientation sexuelle d'une personne physique.

En dehors du cas particulier des recherches dans le domaine de la santé, pour lesquelles des dispositions spécifiques sont prévues (article 44.3 de la loi Informatique et Libertés), de telles données ne pourront être traitées que si l'une des conditions suivantes est remplie :

- si elles ont été **rendues manifestement publiques par la personne concernée** ;
- si le traitement est nécessaire à la recherche publique au sens du code de la recherche, et que ces utilisations sont rendues nécessaires « pour des motifs d'intérêt public important », après avis motivé et publié de la CNIL qu'il leur appartient donc de solliciter au titre de l'article 44.6 de la loi Informatique et Libertés.

  La recherche privée, qui ne rentre pas dans les critères définis dans le code de la recherche, ne peut donc pas mobiliser cette exception prévue par la loi Informatique et Libertés.
- si un texte de droit national ou européen (par exemple un décret en Conseil d'État pris après avis de la CNIL) autorise ce traitement au titre d'un **motif d'intérêt public important**.
- **ou si les personnes concernées y ont expressément consenti**.

Focus

Comment apprécier si des données ont été manifestement rendues publiques : le cas des réseaux sociaux

Dans un arrêt du 4 juillet 2023 (affaire C-252/21), la CJUE a souligné le fait que cette dérogation doit être interprétée de manière restrictive. Elle ne s'applique pas aux données concernant d'autres personnes que celles les ayant rendues publiques et il importe de vérifier si la personne concernée a entendu, de manière explicite et par un acte positif clair, rendre accessibles au grand public les données en question.

[18] La CNIL rappelle qu'une photographie ou un vidéogramme ne constituent pas en eux-mêmes des données biométriques. En revanche, constituerait un tel traitement la création de gabarits biométriques (c'est-à-dire de représentations numériques des caractéristiques physiques des personnes comme des visages), à partir de photographies collectées sur Internet pour fournir un moteur de recherche permettant d'identifier un individu à partir d'une image (voir la délibération de la formation restreinte de la CNIL SAN-2022-019 du 17 octobre 2022)

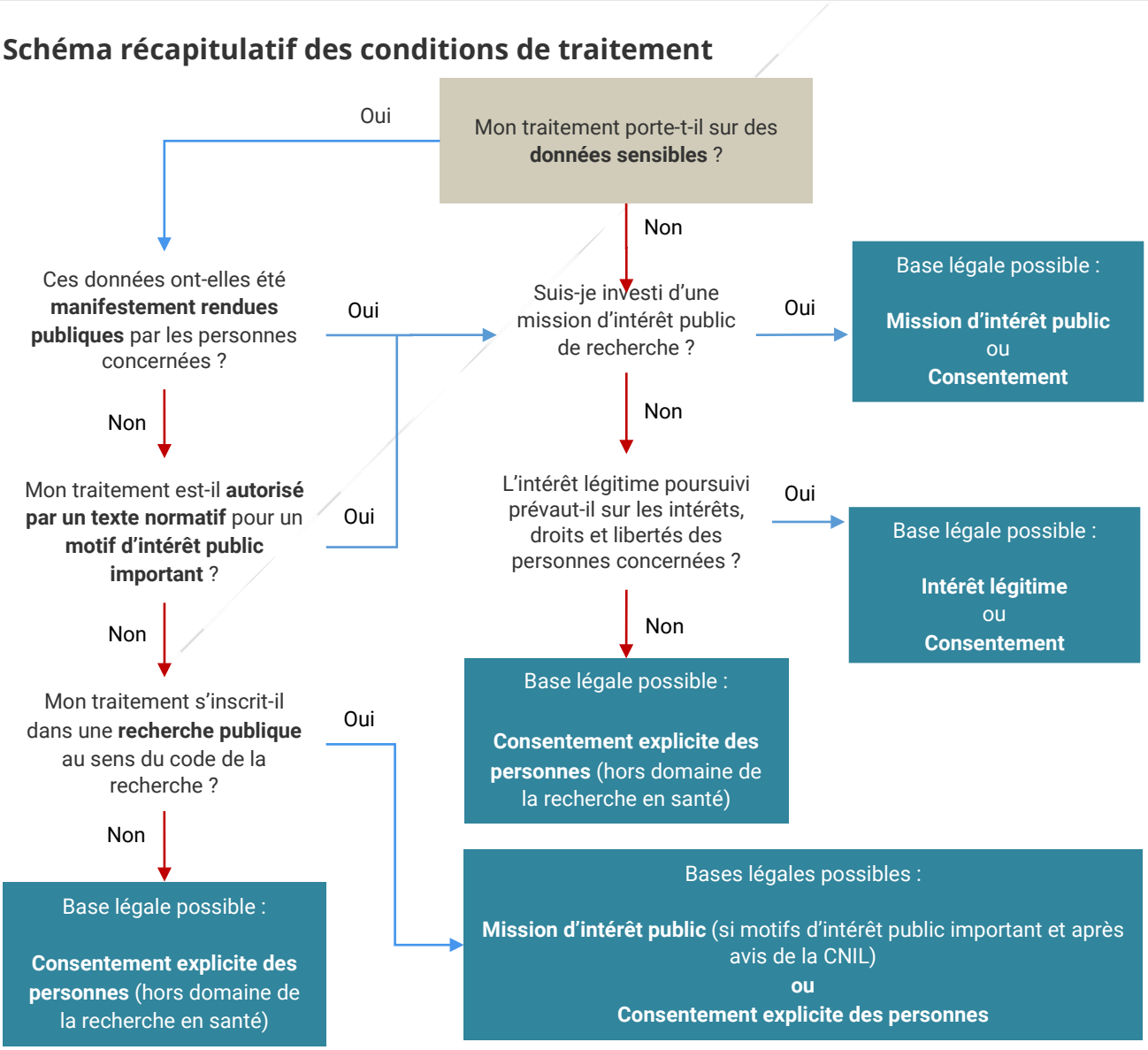
Ainsi, lorsqu'un internaute insère des données sur des sites Internet ou applications en rapport avec les catégories de données sensibles visées à l'article 9 du RGPD, ou active des boutons de sélection intégrés à ces sites et à ces applications (ex. : boutons « j'aime » ou « partager ») ou les boutons lui permettant de s'identifier en utilisant les identifiants de connexion liés à son compte d'utilisateur du réseau social, son numéro de téléphone ou son adresse électronique, il ne peut être considéré comme ayant rendues manifestement publiques les données ainsi insérées ou résultant de l'activation de ces boutons que dans le cas où il a explicitement exprimé son choix au préalable, le cas échéant sur la base d'un paramétrage individuel effectué en toute connaissance de cause, de rendre les données le concernant publiquement accessibles à un nombre illimité de personnes.

À noter : les données manifestement rendues publiques ne se limitent pas aux données accessibles sur les réseaux sociaux.

Exemple

Des chercheurs au sein d'une université publique entreprennent une recherche dans le domaine des sciences politiques. A cet égard, ils traitent des données publiquement disponibles sur des sites institutionnels afin d'étudier les conditions d'exécution par les élus de leur mandat politique (taux de présence, nombre de rapports, sens de leurs votes sur différentes thématiques, etc.). Dans la mesure où le responsable de traitement est une université publique et que les orientations politiques sont manifestement rendues publiques par les personnes concernées dans le cadre de leur activité, le traitement de ces données sensibles est licite.

Schéma récapitulatif des conditions de traitement



À noter

En dehors des données sensibles visées par l'article 9 du RGPD, d'autres catégories de données font l'objet de mesures de protection spécifiques comme :

- **les données relatives à des infractions et condamnations**, qui ne peuvent être légalement traitées que par certaines catégories d'acteurs (autorités publiques et les personnes morales gérant un service public, agissant dans le cadre de leurs attributions légales, associations d'aide aux victimes conventionnées par le ministère de la justice, etc.) ;
- **ou encore celles relatives à des mineurs**, massivement présents en ligne et auxquels le cadre juridique de la protection des données accorde une attention particulière, s'agissant notamment des conditions de prise en compte de leurs droits.

Les mesures de minimisation des données et la proportionnalité des réutilisations

S'agissant des traitements à des fins de recherche scientifique, l'article 89 du RGPD insiste sur l'importance de la mise en œuvre de mesures permettant d'assurer le principe de minimisation des données.

Ne collecter que les données nécessaires à l'objectif poursuivi

Tout d'abord, les réutilisateurs de données doivent **mener une réflexion sur le champ des données « strictement nécessaires » à leur recherche en veillant à ne pas collecter pas davantage de données que nécessaires**.

Lorsque la collecte de données n'est pas réalisée manuellement mais au moyen de procédés automatisés, cela implique :

- **de privilégier autant que possible [19] l'utilisation d'interfaces de programmation applicative (« API »)[20] permettant de cibler les données recueillies** (cf. recommandation de la CNIL n° 2023-050 du 25 mai 2023),
- **de ne recourir à des techniques de moissonnage de données (*scraping*) que de manière ciblée et pertinente**, c'est-à-dire :
  - en définissant **des critères précis et pertinents de collecte** afin d'opérer une distinction quant à la nature des données collectées ;
  - **en limitant le risque de recueil de données sensibles**, lors de la sélection des mots clés ou catégories de contenus, par exemple ;
- de **supprimer toute donnée non pertinente, en particulier s'il s'agit d'une donnée sensible au sens de la réglementation, dès qu'elle est identifiée comme telle**. Cela est valable quelle que soit le mode de collecte, qu'il s'agisse du téléchargement d'un fichier librement

À noter

Si le recours à des techniques de moissonnage n'est pas en soi incompatible avec les exigences du RGPD, il est susceptible de compromettre le bon fonctionnement des sites « aspirés » (par exemple en occasionnant des dénis de service si trop de requêtes sont adressées simultanément), ou d'être interdits par d'autres réglementations (par exemple par des conditions générales d'utilisation qui s'appuieraient sur le droit des producteurs de bases de données ou sur le droit d'auteur). A cet égard, les organismes de recherche peuvent envisager de bénéficier de l'exception de « fouille de texte et de données » en vertu du code de la propriété intellectuelle (art. L122-5 et 122-5-3).

[19] A cet égard, la disponibilité d'API ne devrait pas exclure catégoriquement tout recours à des techniques de moissonnage des données, par exemple lorsqu'il s'agit d'étudier le comportement et/ou les risques systémiques des plateformes au sens de l'article 40 du règlement 2022/2065 sur les services numériques.

[20] Voir la recommandation technique de la CNIL sur le partage de données par API (novembre 2023).

accessible, du recours à des API ou à des techniques de moissonnage de données, dès lors qu'un tri exhaustif n'étant pas toujours possible de manière automatisée.

L'anonymisation ou la pseudonymisation des données collectées

Il est essentiel d'anonymiser ou pseudonymiser les données quand leur exploitation sous une forme identifiante (directement ou non) n'est pas nécessaire.

Toutes les fois où un traitement d'anonymisation ou de pseudonymisation des données collectées en ligne peut être mis en œuvre, il convient d'y recourir. Les enjeux et avantages de ces techniques sont présentées dans une fiche pratique de la CNIL.

De tels traitements, qui excluent ou limitent le risque d'identification des personnes concernées, **doivent être envisagés aussi bien au stade de la collecte des données qu'au stade de leur exploitation**.

Exemples

- **L'anonymisation** peut être très pertinente dans le cadre d'une étude quantitative, lorsque des résultats agrégés suffisent à produire des analyses pertinentes (par exemple : le nombre de personnes s'étant publiquement prononcé sur tel sujet sur un réseau social de type microblogging).
- Lorsque l'anonymisation des données nuirait au déroulement de la recherche, par exemple dans le cadre d'une approche qualitative qui supposerait l'analyse approfondie de contenus publiés par des personnes concernées, la **pseudonymisation** des données et leur conservation sous une forme non directement identifiante, pendant la durée de l'analyse des données collectées peuvent être envisagée (ce fut notamment le cas d'un projet de traitement ayant pour finalité une recherche sur les impacts pour la vie privée des publications d'informations librement accessibles sur les réseaux sociaux, cf. délibération de la CNIL n° 2018-151 du 3 mai 2018).
- À l'inverse, le traitement de données brutes directement identifiantes (telles que des profils publiés sur un réseau social), pendant une longue durée nécessitera une analyse approfondie de la base légale, en particulier s'il s'agit de l'intérêt légitime, dès lors que les risques pour les intérêts, droits et libertés seront plus importants en l'absence de toute mesure de pseudonymisation.

Par ailleurs, si les finalités pour lesquelles des données personnelles sont traitées n'imposent pas ou plus au réutilisateur d'identifier les personnes, celui-ci n'est pas tenu de conserver, d'obtenir ou de traiter des informations supplémentaires pour les identifier à la seule fin de respecter le présent règlement, par exemple pour répondre à des demandes d'exercice des droits. En revanche, dans un tel cas, les personnes en sont si possible informées, et peuvent fournir des informations complémentaires permettant de les identifier afin d'exercer leurs droits (article 11 du RGPD).

Documentation de la conformité

La démonstration, documentée par le responsable de traitement du bon respect des principes précédemment évoqués est essentielle.

Cette évaluation sera d'ailleurs nécessaire dans le cadre de la réalisation de l'analyse d'impact relative à la protection des données (AIPD), **obligatoire avant la mise en œuvre du traitement si ce dernier est susceptible d'engendrer un risque élevé pour les personnes**.

Ce sera souvent le cas de l'utilisation d'outils de moissonnage de données, mais aussi plus généralement de toute collecte à grande échelle portant sur des personnes vulnérables, ou encore de croisement de données ou de profilage portant sur des données sensibles ou à caractère hautement personnel.

Cette analyse permet de documenter les mesures lors de la collecte et la constitution d'un jeu de données (par exemple, en détaillant les critères de sélection des données, la mise en œuvre de leur anonymisation ou de leur pseudonymisation, etc.). L'AIPD permettra aussi de s'assurer de la pertinence du choix et du respect d'une durée de conservation limitée des données, de la confidentialité et la sécurité des données. Il s'agit notamment de prendre en compte les risques que poseraient d'éventuelles réutilisations de ces données.

L'information des personnes concernées

Le RGPD prévoit que tout responsable de traitement doit fournir des mentions d'information aux personnes concernées sur le traitement de leurs données. Cette information est en principe directement fournie à chaque personne dont les données sont traitées.

Lorsque les données ne sont pas collectées auprès de la personne concernée comme c'est le cas de la réutilisation des données publiquement accessibles, le RGPD prévoit que la fourniture d'une information publique peut s'avérer suffisante (par exemple sur le site du responsable du traitement), notamment dans les deux cas suivants :

- Plus précisément, **l'article 14.5.b du RGPD prévoit qu'une information générale peut suffire lorsqu'une information individuelle se révèle impossible ou exigerait des efforts disproportionnés, en particulier s'agissant des traitements à des fins de recherche scientifique.**
  En pratique, le caractère proportionné s'apprécie en fonction de l'atteinte portée à la vie privée des personnes dont les données sont traitées et de la difficulté et le coût d'une information individuelle :

  - lorsqu'une information par voie électronique est possible, elle est généralement requise ;

  - lorsque les données disponibles sont des données postales, une analyse au cas par cas est nécessaire, en prenant en compte notamment le coût de l'information, la faiblesse ou l'importance des risques que le traitement peut faire encourir aux personnes, le fait que les personnes peuvent ou non raisonnablement s'attendre au traitement de leurs données ;

  - beaucoup de jeux de données pseudonymisées mis à disposition de chercheurs sur Internet ne contiennent pas de données de contact, conformément au principe de minimisation décrit ci-dessus. Lorsque les données de contact ne sont pas ou plus disponibles, il est généralement plus facile de démontrer qu'une information individuelle exigerait des efforts disproportionnés. Une analyse de proportionnalité est cependant nécessaire pour s'assurer que le défaut d'information individuelle ne remet pas en cause la licéité du traitement.

- La fourniture d'une information générale est aussi envisageable lorsqu'une information individuelle est susceptible de rendre impossible ou de compromettre gravement la réalisation des objectifs dudit traitement de chaque personne concernée.

  Exemple

  Ce pourrait être le cas des travaux de recherche scientifique impliquant l'analyse de données diffusées sur les réseaux sociaux, si l'information des personnes concernées risque de les conduire à modifier leur comportement en ligne et, ainsi, de biaiser les travaux en question.

Cette transparence est essentielle pour permettre aux personnes concernées d'exercer leurs droits sur le traitement de leurs données, tels que le droit d'accès, le droit d'opposition, le droit à l'effacement ou le droit à la rectification de leurs données auprès du réutilisateur. Il est à noter que des dérogations sont permises lorsqu'elles sont nécessaires à l'objectif de recherche.

Par ailleurs, la fourniture de mentions d'information complètes (notamment sur la source des données) permet également aux personnes d'exercer leurs droits auprès du diffuseur de leurs données.

Fiche cas d'usage n°4 : le moissonnage de données publiquement accessibles par des autorités publiques dans le cadre de leurs missions

Certaines autorités publiques ont recours à des techniques de moissonnage de données publiquement accessibles dans le cadre de leurs prérogatives de puissance publique.

Pour ce faire, elles extraient des informations de sites web, tels que les réseaux sociaux, au moyen de procédés informatiques appelés outils de moissonnage, ou de « *web scraping* » en anglais.

Compte tenu du risque de surveillance généralisé inhérent à de telles techniques, et plus généralement des enjeux pour les droits et libertés des citoyens qui ne maîtrisent pas les réutilisations de leurs données accessibles en ligne, ces traitements de données personnelles doivent comprendre un certain nombre de garanties.

La nécessité d'un encadrement juridique

Ni le fait que des données soient publiquement accessibles, ni la création volontaire de profils sur les plateformes en ligne n'emporte, par principe, la possibilité de leur moissonnage ainsi que de leurs réutilisations. Les acteurs qui souhaitent les exploiter ne sont pas exonérés de **l'obligation de collecter ces données de manière licite et loyale**.

Pour qu'une autorité publique puisse mettre en œuvre de tels traitements dans le cadre de ses missions, il est souvent nécessaire que le traitement soit autorisé par un texte de droit français ou européen. Par exemple, cela sera notamment rendu nécessaire par l'article 31 de la loi Informatique et Libertés qui exige une autorisation par arrêté ou par décret (pris après avis motivé et publié de la CNIL) pour les traitements de détection d'infraction mis en œuvre pour le compte de l'Etat par les autorités compétentes (c'était notamment le cas du traitement mis en œuvre par VIGINUM ayant recours à des outils de moissonnage de données visant à détecter et caractériser les opérations d'ingérence numérique étrangères aux fins de manipulation de l'information sur les plateformes en ligne[21]).

Cet encadrement spécifique devra faire apparaître de manière explicite les finalités des traitements.

Lorsque le traitement affecte les droits et libertés fondamentaux dont bénéficie le citoyen au sens de l'article 34 de la constitution (telles que la liberté d'expression et la liberté d'opinion), **seule une loi pourra autoriser sa mise en œuvre en définissant son périmètre et des garanties suffisantes**.

À cet égard, la **limitation de la collecte aux données librement accessibles** (c'est-à-dire aux contenus accessibles à tout utilisateur non inscrit sur le site en question et sans création d'un compte) et **manifestement rendues publiques par la personne concernée** (ce qui exclut notamment les commentaires ou interactions de tiers portant sur la personne concernée) est une garantie essentielle. Elle a notamment été déterminante pour apprécier la proportionnalité de l'expérimentation de l'exploitation de données rendues publiques sur les sites internet des opérateurs de plateformes en ligne par les administrations fiscale et douanière.

Il est possible de citer d'autres garanties ayant déjà été prévues par des textes, telles que :

- la limitation des services sur lesquels le recours aux outils de moissonnage de données est possible (par exemples aux seuls sites internet des opérateurs de plateforme de mise en relation de plusieurs parties en vue de la vente d'un bien, de la fourniture d'un service ou de l'échange ou du partage d'un contenu, d'un bien ou d'un service, dépassant un seuil important de visiteurs) ;
- l'interdiction de prendre une décision entièrement automatisée sur la base d'un traitement de moissonnage de données ;
- l'interdiction de certains procédés, tels que le recours à un système de reconnaissance faciale ou d'identification vocale ;
- l'existence d'un contrôle externe, effectif et adapté aux caractéristiques du traitement envisagé, ou encore
- l'exigence d'une transparence accrue, notamment à travers la production de rapports annuels publics.

[21] Ce traitement est prévu par le décret n° 2021-1587 du 7 décembre 2021 portant autorisation d'un traitement automatisé de données à caractère personnel dans le but d'identifier les ingérences numériques étrangères.

La minimisation de la collecte

Le respect du principe de minimisation est fondamental compte tenu du volume important de données et plus particulièrement, d'informations non nécessaires ou non pertinentes susceptibles d'être moissonnées.

Pour cette raison, **la CNIL recommande d'utiliser une interface de partage de données (API) ou toute autre fonctionnalité de mise à disposition par le site ou la plateforme lorsque cela est possible, plutôt que de recourir à des outils de moissonnage de données.**

Lorsqu'une autorité décide de recourir à des outils de moissonnage de données, elle doit veiller à **définir, en amont de la mise en œuvre du traitement, des critères précis de collecte afin d'opérer une distinction quant à la nature des données collectées**. Cela suppose en amont d'élaborer des indicateurs et critères de pertinence, permettant d'orienter et cibler la collecte.

Il convient ensuite **de s'assurer de la suppression des données non-pertinentes immédiatement après leur collecte, notamment lorsqu'elles sont sensibles** (par exemple des informations sur la santé, la religion ou l'orientation sexuelle des personnes).

Un tri exhaustif n'étant pas toujours possible de manière automatisée, l'autorité publique doit **veiller à supprimer toute donnée non pertinente dès qu'elle est identifiée comme telle** dans le cadre de l'exploitation des données ainsi collectées.

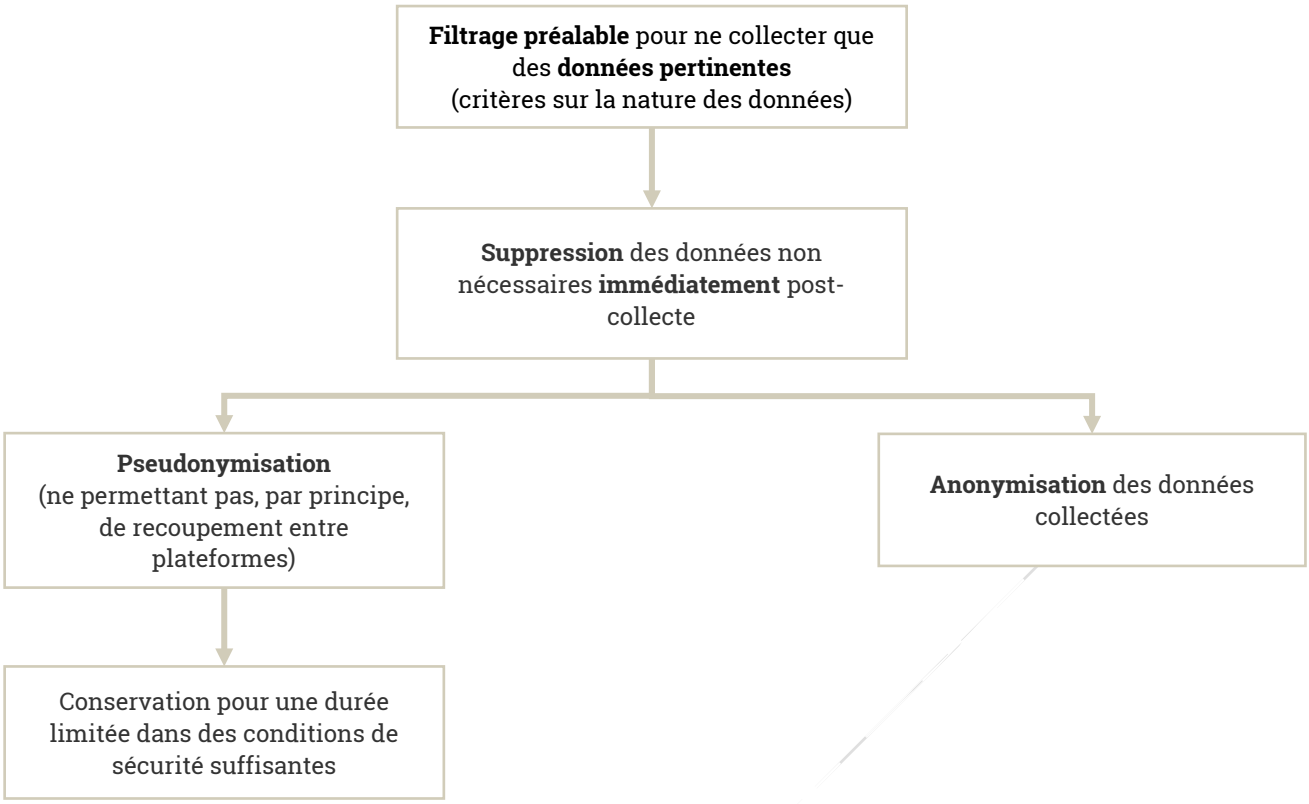
Par exemple, dans le cas de VIGINUM, l'administration doit s'assurer que seules les informations en lien avec les thématiques et acteurs présumés étrangers d'intérêt soient collectées, notamment en utilisant un procédé automatisé de suppression des données non pertinentes (afin de garantir que les données initialement collectées ne seront pas accessibles aux personnes habilitées avant l'opération de suppression automatisée des données non pertinentes).

La CNIL recommande qu'une anonymisation ou une pseudonymisation des données soit effectuée juste après la collecte des données et que le recoupement, sur plusieurs plateformes en ligne, à partir des identifiants collectés ne soit possible qu'après avoir conduit une analyse documentée dans l'AIPD afin de mettre en balance l'intérêt de ce recoupement et les risques pour les droits et libertés des personnes concernées.

La sécurité et la conservation des données

Les données moissonnées doivent être correctement sécurisées au regard des risques encourus. Par exemple, dans le cas des expérimentations mises en œuvre par le PEReN (pôle d'expertise de la régulation numérique de traitements), une habilitation individualisée des accès aux données et une journalisation de ces accès est prévue, La CNIL a recommandé en complément la mise en place d'un mécanisme de double authentification.

Ces données ne doivent pas être conservées plus longtemps que nécessaire au regard de la finalité de leur collecte. La mise en œuvre de ces traitements ne doit pas non plus engendrer une collecte permanente de données sur les plateformes concernées. La CNIL recommande de définir au cas par cas la durée de la collecte et la période de récupération des données (profondeur historique), qui devra également être documentée. À titre d'exemple, le texte encadrant le traitement de VIGINUM a précisé la période au cours de laquelle les données peuvent être collectées, soit une période maximale de sept jours.



Lorsque l'autorité a recours aux services d'un sous-traitant pour réaliser de telles opérations, un contrat doit être conclu, contenant toutes les mentions prévues par l'article 28 du RGPD (telles que celles relatives aux instructions documentées et aux mesures de sécurité). Par ailleurs, le responsable de traitement doit s'assurer que son sous-traitant présente des garanties suffisantes pour la conformité des traitements (pour en savoir plus : voir le guide du sous-traitant).

L'information des personnes et l'exercice de leurs droits

En principe, l'autorité doit informer les personnes concernées du traitement qu'elle met en œuvre.

Une information générale peut s'avérer suffisante, par exemple sur le site internet de l'autorité en question lorsque la fourniture d'une information individuelle exigerait des efforts disproportionnés pour l'autorité ou qu'elle est écartée par le texte autorisant le traitement.

Cette information permet notamment aux personnes concernées d'exercer leurs droits sur leurs données auprès de l'autorité. Il peut s'agir du droit d'accès, du droit à la rectification ou à l'effacement des données.

Lorsque l'autorité reçoit une demande d'exercice de l'un de ces droits après avoir procédé à l'anonymisation ou la pseudonymisation des données, il est possible pour l'autorité de démontrer qu'elle n'est pas en mesure d'identifier la personne concernée. Cependant, les personnes devraient toujours pouvoir fournir des informations complémentaires permettant de les identifier.

Enfin, le droit d'opposition devra parfois être envisagé pour certains traitements. L'effectivité de ce droit d'opposition pourra notamment être garantie au moyen d'une liste repoussoir, y compris en amont du traitement.
S'agissant de la liste repoussoir mise en œuvre par le PEReN dans le cadre de ses expérimentations, la CNIL a par exemple recommandé de ne conserver que des empreintes cryptographiques des identifiants des personnes concernées (en supprimant tout lien entre les identifiants d'une même personne) et de supprimer les données brutes immédiatement. Elle a aussi recommandé de permettre aux personnes concernées de choisir la durée de prise en compte de leur opposition.

La réalisation d'une AIPD

La démonstration, documentée par le responsable de traitement du bon respect des principes précédemment évoqués est essentielle avant tout déploiement de tels dispositifs.

Cette évaluation sera d'ailleurs nécessaire dans le cadre de la réalisation de l'analyse d'impact relative à la protection des données (AIPD), **obligatoire avant la mise en œuvre du traitement si ce dernier est susceptible d'engendrer un risque élevé pour les personnes**. Ce sera souvent le cas de l'utilisation d'outils de moissonnage de données, dès lors qu'il s'agit d'un usage innovant conduisant à une collecte à grande échelle, portant souvent sur des données sensibles ou à caractère hautement personnel et pouvant conduire à une décision automatisée avec effet légal ou similaire.

Si le traitement est mis en œuvre par une autorité compétente à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière, pour le compte de l'État, cette analyse d'impact doit même être adressée à la CNIL dans le cadre d'une demande d'avis.