Comment procéder à l’exploitation des données en conformité avec la réglementation ?

# Introduction :

**Qu’est-ce que le RGPD ?**

C’est un règlement européen entré en vigueur en **2016,** qui a pour but de renforcer **le contrôle par les citoyens de l’utilisation de leurs données personnelles. Ce règlement offre aussi** un cadre juridique précis permettant de développer des activités numériques au sein de l’Union européenne en considérant acquise la confiance de leurs utilisateurs.

**A qui s’applique le RGPD ?**

Le RGPD s’applique à toute **organisation**, qu’elle soit **publique ou privée**, peu importe sa taille, son activité et son secteur d’exploitation), des lors qu’elle **traite des données personnelles** pour son compte ou en tant que sous-traitant et qu’elle est établie sur le territoire de l’Union Européenne ou que son activité touche directement des citoyens européens.

Dès lors la question qui reste en suspens est :

**Qu’est-ce qu’une donnée personnelle ?**

C’est, je cite "toute information se rapportant à une personne physique identifiée ou identifiable". Pour être plus démonstratif, vis-à-vis d’une personne, cela peut être : nom, prénom, pseudonyme, date de naissance, photos, enregistrements sonores de voix, numéro de plaque d’immatriculation. L’identification peut se faire directement ou bien par croisement de données. Exemple : la plaque d’immatriculation peut comporter plusieurs utilisateurs donc le croisement avec le nom ou la date de naissance permet de distinguer formellement une personne.

Pour finir, **qu’est-ce que le traitement de ces fameuses données personnelles ?**

Une action peut être catégorisée comme traitement des données personnelles dès lors qu’elle manipule des données personnelles, quel que soit le procédé utilisé (collecte, enregistrement, organisation, conservation, adaptation, modification, extraction, consultation, utilisation).

1. Les principales clés permettant le développement d'applications ou de sites web respectueux du RGPD

I.1) Sensibiliser et préparer le développement

a) Le DPO, délégué à la protection des données ou Chief Data Officer

Le grand principe permettant le développement d'applications ou de sites web respectueux du RGPD est de travailler de façon alerte. Tout au long du développement d’un projet que ce soit en autonomie ou au sein d’une équipe, il ne faut jamais perdre de vue la protection des données.

Le DPO incarne le référent pour toute question des collaborateurs et dicte la bonne conduite à adopter vis-à-vis de la prise en compte du RGPD au cours du développement. Sa mission est de s’assurer de la conformité du développement vis-à-vis du règlement européen.

La nomination d’un DPO est obligatoire pour les organismes publics et les entreprises qui mène des traitements à grandes échelles ou sur des données sensibles.

Il est toutefois conseiller d’élire un DPO car il permet à tous de comprendre et respecter les obligations du règlement. De plus il sert d’interlocuteur avec les autorités de protection des données et minimise les risques de litiges.

b) Méthodologie

Le DPO peut proposer des méthodologies de développement afin de le centrer sur la sécurité des données et de la vie privée.

Considérons rapidement l’opposition entre le Privacy by Design et le Privacy By Default :

Ce premier prend en compte la protection de la vie privée dès la conception du produit ou du service. Etant une préoccupation antérieure au développement du produit, le respect de la vie privée en fait partie intégrante. Cela suppose une approche réfléchie et évolutive de la protection des données au cours du développement.

De l’autre côté, le privacy By Default, prévoit la mise en place du plus haut degré de protection des données personnelles et l’activation de ces processus par défaut, ie, sans configuration préalable de la part de l’utilisateur.

1.2) S’Informer et se conformer aux standards

a) Le registre des activités de traitements.

Le registre des activités de traitements est un document qui vise à centraliser les informations concernant les traitements induits par l’application, les données qui leur sont nécessaires et leur criticité. Tenir ce registre permet donc de faciliter la restitution des preuves qui permettent de démontrer que le développement est à tout instant en conformité avec le RGDP. La tenue du registre d’activité s’intègre parfaitement dans une méthodologie de Privacy By Design. Le registre d’activité attire l’attention notamment sur : la nécessité, les types de données collectées et traitées par votre programme, la base légale sur laquelle se fonde vos traitements de données, les clauses contractuelles liant l’entreprise à ses sous-traitants, les modalités d'exercice des droits, les mesures mises en place pour sécuriser vos traitements.

Ainsi dans le registre d’activité, il faut créer une fiche pour chaque activité recensée en précisant :

* La finalité, i.e., ce pour quoi la donnée va être collectée.
* Le type des données utilisées.
* Le bénéficiaire ou destinataire qui va avoir accès aux données.
* La durée de conservation de ces données.

La constitution de ce registre à un second avantage du point de vue de l’entreprise, elle lui permet de s’interroger sur les données dont elle a réellement besoin.

* Ce traitement nécessite-t-il réellement l’ensemble des données que je m’apprête à collecter ?
* L’un des traitements manipule-t-il des données dites sensible ? Ai-je vérifié que j’ai les autorisations correspondantes ?
* Ai-je réalisé une bonne gestion des accès afin de permettre d’accorder un accès seulement aux données nécessaires ?
* La durée de conservation des données est-elle adaptée ?
* Les données collectées sont-elles protégées au regard de leur criticité ?

Sa création et sa mise à jour sont ainsi l’occasion de faire le tri dans les données collectées et permet de redéfinir un plan d’action de mise en conformité des traitements aux règles de protection des données.

Minimiser les données collectées : <https://www.cnil.fr/fr/minimiser-les-donnees-collectees>

Les durées de conservations : <https://www.cnil.fr/fr/les-durees-de-conservation-des-donnees>

I.2) S’informer et se conformer aux standards

b) La gestion des risques induits par les traitements, l’AIPD.

Chaque traitement sur des données à caractère personnelle peut engendrer des risques de divulgation involontaire des données utilisateurs. Ainsi dans le but de couvrir ces potentiels risques, il est nécessaire de les analyser pour connaitre leur criticité.

Pour cela, la CNIL préconise le recours à une Analyse d’impact relative à la protection des Données (AIPD). Afin de la réalisée en bonne et due forme, il est possible de se référer aux méthodes, modèles de documents et au catalogue des bonnes pratiques accessible sur le site de la CNIL. Elle fournie également un outil de conception de cette AIPD sur son site.

<https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-1-fr-methode.pdf> <https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-2-fr-modeles.pdf> <https://www.cnil.fr/fr/outil-pia-telechargez-et-installez-le-logiciel-de-la-cnil> <https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-3-fr-basesdeconnaissances.pdf>

Pour certains traitements l’AIPD est obligatoire, pour d’autres non. Il est possible de se référer aux listes et à ce chemin logique :

<https://www.cnil.fr/sites/default/files/atoms/files/infographie_aipd.pdf>

<https://www.cnil.fr/sites/default/files/atoms/files/liste-traitements-aipd-requise.pdf>

<https://www.cnil.fr/sites/default/files/atoms/files/liste-traitements-aipd-non-requise.pdf>

Mais alors qu’est-ce qu’une analyse d'impact relative à la protection des données (AIPD) et quand doit-elle être mise en place me direz-vous ?

C’est un outil d’évaluation du potentiel impact du traitement sur la vie privée dès lors qu’il est susceptible d’impliquer des risques conséquents sur les droits et libertés des utilisateurs. Elle doit donc être réalisée avant la mise en œuvre du traitement.

I.3) Informer et protéger

a) Les équipes et/ou sous-traitants.

Le rôle du DPO, est comme nous l’avons dit, d’informer et de tenir à jour les équipes des directives à suivre. Cependant les traitements ne sont pas toujours réalisés en interne et il se peut que l’on ait recours à des sous-traitants. Mais alors quelles sont leurs obligations vis-à-vis de l’entreprise qui fournies les données et de ses utilisateurs. La réglementation prévoit que si un organisme traite ou collecte des données pour un tier, il a des devoirs notamment en matière de sécurité et de protection des données qui lui sont attribuées, de confidentialité, de documentation de ses activités pour le compte de son client. Les sous-traitants, de part leur statut, ont une obligation de conseil et d’aide de mise en conformité au RGPD dès lors qu’ils s’aperçoivent d’un manquement. Ils sont priés de tenir un registre des activités et d’aider à la mise en œuvre d’une AIPD.

Finalement, tout cela doit être concilier dans un contrat entre les deux parties.

I.3) Informer et protéger

b) Les utilisateurs

Quel que soit l’organisme qui traite des données personnelles, ses utilisateurs possèdent des droits qui ne peuvent leur être refuser. Ainsi les organismes collecteurs sont assujettis à une obligation de transparence qui leur impose d’informer les personnes concernées par la collecte de leurs données personnelles de leurs droits. On dénombre parmi ces droits : les droits d’accès, de rectification, de limitation, d’opposition et d’effacement.

i. Comment tenir informer les utilisateurs sur leurs droits

Il est nécessaire de tenir informer l’utilisateur de ses droits lors de la collecte de ses données faite directement par un formulaire, lors d’achats en ligne ou indirectement comme des données de géolocalisation ou de trackings. Il est nécessaire de réinformer l’utilisateur de ses droits dès lors que les modalités d’utilisation de ses données sont amenées à changer. Par exemple : nouvelle finalité ou bien nouveau destinataire.

En plus des droits des utilisateurs les organismes de collecte doivent fournir : L'identité et les coordonnées de l’organisme qui collecte les données, les finalités, la base légale sur laquelle repose le traitement des données, le caractère obligatoire ou facultatif du recueil des données (ce qui suppose une réflexion en amont sur l’utilité de collecter ces données au vu de l’objectif poursuivi – principe de « minimisation » des données) et les conséquences pour la personne en cas de non-fourniture des données, les destinataires des données, la durée de conservation des données, les coordonnées du délégué à la protection des données de l’organisme, le droit d’introduire une réclamation auprès de la CNIL.

<https://www.cnil.fr/fr/informer-les-personnes>

ii. Comment permettre aux utilisateurs d’exercer leurs droits

<https://www.cnil.fr/fr/preparer-lexercice-des-droits-des-personnes>

Il faut savoir que dès lors qu’une organisation collecte et traite des données, elle doit fournir des explications à ses utilisateurs sur où et comment exercer leurs droits. Un moyen simple est d’inclure les coordonnées du DPO à contacter lors des étapes d’informations des utilisateurs. Une autre manière est d’implémenter directement au sein de l’application des processus d’exercice des droits.

Attention toutefois, ces droits d’accès, de modifications doivent être accessibles mais sécurisés de sorte qu’une personne tiers ne puisse y avoir accès.

Prenons un exemple :

Droit à l'effacement : les personnes peuvent demander l’effacement de la totalité des données que l’entreprise détient à leur sujet.

On peut imaginer de prévoir une option dans les comptes utilisateurs permettant d’effacer toutes les données relatives à une personne après vérifications du mot de passe.

Attention il faut veiller également à soumettre une notification aux sous-traitants afin qu’ils effacent eux aussi les données relatives à cette personne ainsi qu’à effacer les données des sauvegardes dans le but de rendre toute restauration des données impossible.

Pour récapituler :

L’objectif de la soumission à l’ensemble des préconisations de la CNIL est de pouvoir à tout moment du développement et du lancement de l’application de justifier sa conformité au RGPD. Cela implique notamment un réexamen et une actualisation réguliers de la documentation des développements afin qu'elle soit en permanence en cohérence avec les fonctionnalités.

Pour prouver sa conformité au règlement, il faut constituer et regrouper la documentation suivante:

* La documentation sur les traitements de données personnelles tels que : registre de traitement, AIPD, encadrement des transferts de données hors de l’Union européenne ;
* La documentation concernant l’information des personnes telle que : les mentions d’information et les procédures mises en place pour l’exercice des droits ;
* La documentation concernant les contrats définissant les rôles et les responsabilités des différents acteurs tels que : contrats avec les sous-traitants, preuves du consentement des personnes concernées …

1. Les points spécifiques à souligner

II.1) La sécurisation des données

1. Sécurisé son environnement de développement

Conserver la sécurité des serveurs de production, de développement, d'intégration continue ainsi que les postes de travail des développeurs joue un rôle primordial dans la protection des données car ils en constituent un accès à un grand nombre.

La première étape est donc d’évaluer les risques inhérents aux outils et processus utilisés pour le développement. Attention aux accès sur les outils dans le cloud, typiquement GitHub. Il est nécessaire de recenser les mesures de sécurité existantes et de définir une personne en charge de la couverture de risques.

La seconde étape est de sécuriser les serveurs et les postes de travail d'une façon homogène et reproductible afin que toute nouvelle personne puisse se mettre à jour vis-à-vis des dispositions liées à la sécurité des données. La CNIL offre une liste de recommandation sur ce sujet. Mettre en commun dans un document les différentes mesures et configurations peut être un bon moyen de les homogénéisé et de les rendre reproductibles. Enfin les mises à jour régulières sont indispensables à l’efficacité du système de sécurisation des données.

<https://www.cnil.fr/sites/default/files/atoms/files/cnil_guide_securite_personnelle.pdf>

Une dernière étape peut être d’optimiser la gestion des accès aux données et de tracer toute intervention sur les jeux de données tout cela répertorié dans des journaux de logs.

1. La gestion du code source

De même que l’environnement de travail, la gestion du code source fait partie des actions de protections des données. Il est indispensable de connaitre le contenu du code source car celui-ci peut contenir des données personnelles qu’il ne faudrait divulguer. Pour cela des systèmes de Scanner de code permette de vérifier sa bonne qualité.

De plus, il peut être utile d’avoir recours à un logiciel de versionning, comme Git, qui permet de garder trace des avancées du développement en cours et de naviguer entre les différents stades d’avancement du projet. Il est donc possible d’intervenir sur un bout de code dont la non-conformité au RGPD avait échappé à l’attention du responsable et de le modifier de façon post-dev. Toutefois, utiliser ce genre de logiciel implique des responsabilités notamment en termes d’accès et d’authentification. La CNIL conseille la mise en place d’une authentification forte voire même par clé SSH et que les accès soient nivelés et que les groupes formés aient des permissions spécifiques à leur besoin.

Enfin avant de publier le code en ligne, il faut s’assurer que cette action ne constitue pas un risque pour l’utilisateur. Pour cela des étapes de chiffrements déchiffrements des fichiers sources avant l’envoie du dépôt ainsi qu’une purge de ce dernier une fois le dépôt réalisé peuvent être mise en place.

1. Choix de l’architecture

En s’appuyant sur le registre d’activité et sur l’analyse d’impact, il est possible de conceptualiser l’architecture de l’application en prenant compte du parcours et du cycle de vie des données personnelles en son sein. Le choix des supports de données (stockage local, serveur, cloud) est une étape fondamentale de la conceptualisation de l’application.

La création de schéma de flux de données au sein de l’application peut permettre d’identifier les faiblesses de sécurité. De même lors des flux de données vers les espaces de stockages, il faut veiller à ce que l’intégrité de ces mêmes données soit conserver en route.

Une autre possibilité est d’avoir recours à un prestataire pour l’hébergement. Pour cela, il faut s’assurer que les politiques du prestataire en matière de sécurisation des données et de transparence vis-à-vis de l’utilisateur sont en accord avec celles de l’entreprise et que les serveurs sont localisés dans des zones où le RGPD est en vigueur.

1. La gestion des permissions

La gestion des permissions consista à définir différents profils aux autorisations d’accès et aux habilitations différentes de sorte que chacun n’est accès qu’aux données dont il a réellement besoin. Tout d’abord chaque collaborateur doit être identifiable par un numéro unique ce qui permet l’authentification à un groupe et l’attribution des droits correspondant. Toutes ces connexions peuvent être répertoriées grâce à un système de journalisation (journaux de logs précédemment mentionnés). Ces journaux de logs peuvent présenter l’avantage de facilité la détection de faille de sécurité au sein de l’application ou bien lors de son développement. Des tests d’intrusions peuvent également être menés dans le but de tester la robustesse du système et la visualisation de cette intrusion dans les fichiers de logs.

1. Sécurisation des données

Il est tout aussi important d’être transparent vis-à-vis des utilisateurs dont on collecte les données que d’être le plus opaque possible vis-à-vis des menaces externes. Il ne faut pas minimiser l’impact sur les personnes concernées de la perte, la modification involontaire ou la divulgation de leurs données.

Les mesures à considérer diffèrent selon la criticité des données traitées et des risques qui pèsent sur les personnes en cas d’incident. Ainsi on peut citer des actions comme les mises à jour des différents logiciels et antivirus, le changement régulier des mots de passes ainsi que l’option du chiffrement des données.

Encore une fois La CNIL met à disposition des méthodes et outils pour se conformer à l’un des trois niveaux de sécurité exposés :

Niveau 1 : le minimum pour démarrer

Le niveau de sécurité minimum est dicté par une série de règles essentiels publiés par l’ANSSI et la CPME. Parmi les 12 figurent notamment :

* Le choix de mots de passe de complexité élevée
* Des sauvegardes et des mises à jour régulières
* Séparer la messagerie professionnelle de la messagerie personnelle

(<https://www.cnil.fr/fr/securite-des-donnees-les-regles-essentielles-pour-demarrer>)

Niveau 2 : les mesures d’hygiène pour protéger votre SI

Le second niveau de protection intègre des mesures d'« hygiène informatique ». La CNIL et l’ANSSI disposent de guides en ligne pour aider à mettre en œuvre cette stratégie. Ces mesures de sécurité vis notamment à couvrir les risques :

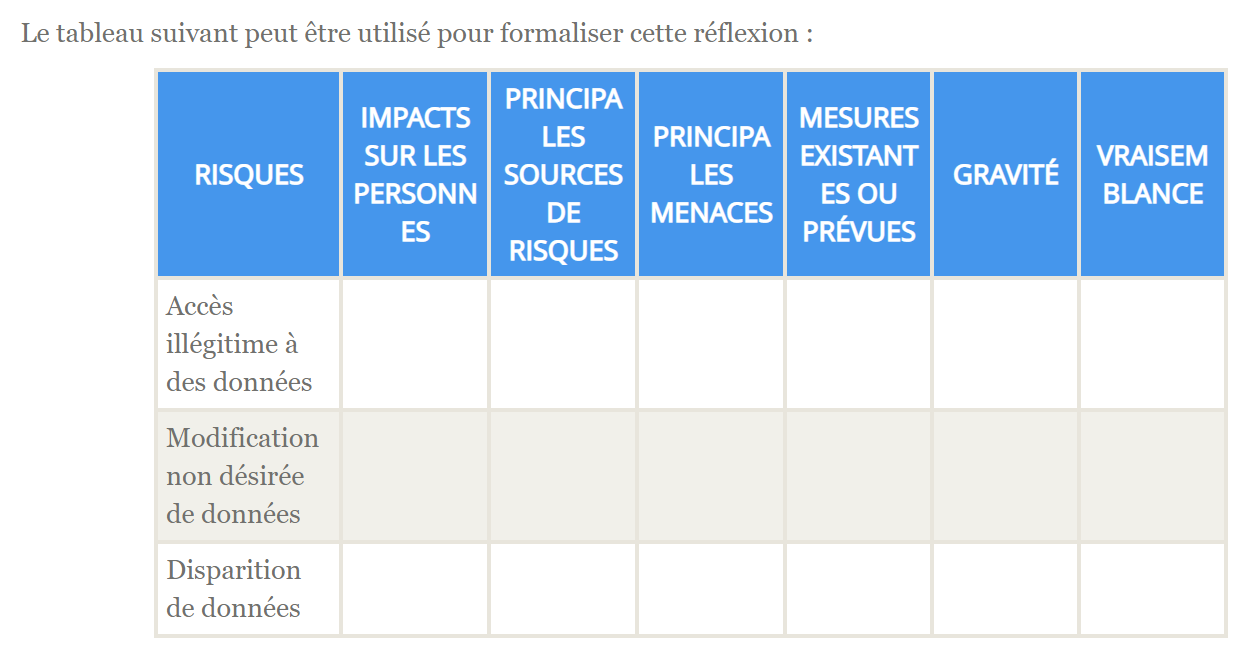
* D’accès non autorisés
* De modifications non désirées
* De disparitions de données

(<https://www.cnil.fr/fr/securite-des-donnees-les-mesures-dhygiene-pour-proteger-votre-systeme-dinformation>)

Niveau 3 : protéger le plus sensible de façon spécifique

Une fois les dispositions antérieures appliquées, il est maintenant nécessaire de satisfaire à l’obligation de sécurité des données traitées en engageant des mesures proportionnelles au risque encouru en cas de défaillance du système de sécurité.

Pour cela une méthode est détaillée par la CNIL :



(<https://www.cnil.fr/fr/securite-des-donnees-proteger-le-plus-sensible-de-maniere-specifique>)

(<https://www.cnil.fr/fr/principes-cles/guide-de-la-securite-des-donnees-personnelles>)

Toute fois, si malgré les dispositifs mis en place, un cas de violation de données est répertorier il faut le signaler à la CNIL dans les 72 heures si cette violation est susceptible de représenter un risque pour les droits et libertés des personnes concernées.

II.2) L’anonymisation des données

Enfin il est possible de rendre les données anonymes, ie, rendre impossible l’identification d’une personne à partir d’un jeu de données et donc de respecter la vie privée des utilisateurs.

**Qu’est-ce que l’anonymisation ?**

L’anonymisation est un traitement qui consiste à utiliser un ensemble de techniques de manière à rendre impossible, en pratique, toute identification de la personne par quelque moyen que ce soit et de manière irréversible.

L’anonymisation ne doit pas être confondue avec la pseudonymisation car malgré la possibilité de traité des données sans identifier de façon direct le propriétaire, ce procédé est dans la plupart du temps réversible ou du moins il est possible d’identifier le propriétaire grâce à un croisement avec un jeu de données externes.

**Pourquoi anonymiser des données personnelles ?**

L’anonymisation n’est pas une obligation vis-à-vis du règlement européen. Il s’agit d’une proposition de solution afin de pouvoir exploiter des données personnelles tout en respectant les droits et libertés des personnes. L’anonymisation des données permet l’exploitation, la réutilisation et le partage de ces dernières dans différentes applications.

**Comment anonymiser tout en préservant au maximum l’utilité du jeu de données ?**

D’un point de vue protection des données l’anonymisation semble être très adaptée mais d’un point de vue affaire on peut se demander comment préserver l’utilité du jeu de données.

Ainsi pour construire un processus d’anonymisation pertinent qui ne nuirai pas à l’exploitation future, il est ainsi conseillé :

* D’identifier les informations à conserver selon leur pertinence.
* De supprimer les éléments d’identification directe ainsi que les valeurs rares qui pourraient permettre un ré-identification aisée des personnes (par exemple, la présence de l’âge des individus peut permettre de ré-identifier très facilement les personnes centenaires)
* De distinguer les informations importantes des informations secondaires ou inutiles (c’est-à-dire supprimables)
* De définir la finesse idéale et acceptable pour chaque information conservée. Ce pré-requis permet de déterminer le procédé d’anonymisation à appliquer, c’est-à-dire l’enchaînement des techniques d’anonymisation à mettre en place. Celles-ci peuvent être regroupées en deux familles : la randomisation et la généralisation.

Comment vérifier l’efficacité de l’anonymisation ?

Les autorités de protection des données européennes définissent trois critères qui permettent de vérifier que le jeu de données est bel et bien anonymisé.

L’individualisation : Une personne en particulier ne doit pas pouvoir être identifiée.

Exemple : une base de données de CV où seuls les nom et prénoms d’une personne auront été remplacés par un numéro (qui ne correspond qu’à elle) permet d’individualiser cette personne. Dans ce cas, cette base de données est considérée comme pseudonymisée et non comme anonymisée.

La corrélation : On ne doit pas pouvoir relier des données de jeu différents par le fait qu’elles concernent la même personne.

Exemple : une base de données cartographique renseignant les adresses de domiciles de particuliers ne peut être considérée comme anonyme si d’autres bases de données, existantes par ailleurs, contiennent ces mêmes adresses avec d’autres données permettant d’identifier les individus.

L’inférence : On ne doit pas pouvoir déduire des informations supplémentaires sur des personnes identifiées dans la base de données.

Exemple : si un jeu de données supposément anonyme contient des informations sur le montant des impôts de personnes ayant répondu à un questionnaire, que tous les hommes ayant entre 20 et 25 ans qui ont répondu sont non imposables, il sera possible de déduire, si on sait que M. X, homme âgé de 24 ans, a répondu au questionnaire, que ce dernier est non imposable.

Sources :

<https://www.cnil.fr/fr/minimiser-les-donnees-collectees>

<https://www.cnil.fr/fr/les-durees-de-conservation-des-donnees>

<https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-1-fr-methode.pdf> <https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-2-fr-modeles.pdf> <https://www.cnil.fr/fr/outil-pia-telechargez-et-installez-le-logiciel-de-la-cnil> <https://www.cnil.fr/sites/default/files/atoms/files/cnil-pia-3-fr-basesdeconnaissances.pdf>

<https://www.cnil.fr/sites/default/files/atoms/files/infographie_aipd.pdf>

<https://www.cnil.fr/sites/default/files/atoms/files/liste-traitements-aipd-requise.pdf>

<https://www.cnil.fr/sites/default/files/atoms/files/liste-traitements-aipd-non-requise.pdf>

<https://www.cnil.fr/fr/informer-les-personnes>

<https://www.cnil.fr/fr/preparer-lexercice-des-droits-des-personnes>

(<https://www.cnil.fr/sites/default/files/atoms/files/formulaire_de_demande_labels-gouvernance-re.docx>)

<https://www.cnil.fr/sites/default/files/atoms/files/cnil_guide_securite_personnelle.pdf>

(<https://www.cnil.fr/fr/securite-des-donnees-les-regles-essentielles-pour-demarrer>)

(<https://www.cnil.fr/fr/securite-des-donnees-les-mesures-dhygiene-pour-proteger-votre-systeme-dinformation>)

(<https://www.cnil.fr/fr/securite-des-donnees-proteger-le-plus-sensible-de-maniere-specifique>)

(<https://www.cnil.fr/fr/principes-cles/guide-de-la-securite-des-donnees-personnelles>)