



# SÉMINAIRE : DONNÉES DE LA RECHERCHE NATURE, CULTURES & VALEURS

---

Association ARISTOTE

18/10/2023

# L'IMPORTANCE DES DONNÉES

Selon le rapport Statista Digital Economy Compass\*, **trente-trois zettaoctets de données numériques** ont été créés dans le monde en 2019.

\*T. Gaudiaut, « La totalité des données créées dans le monde », Statista Digital Economy Compass 2019 in Statista.com, 24 avril 2019

## Quelles différences entre la donnée, la donnée personnelle et l'information ?

### Information

- Protégée par le secret d'affaires (L. 151-1s. ccom)

### Donnée brute

- Protégée par la propriété intellectuelle

### Donnée personnelle

- Protégée par le RGPD

# LES DONNÉES DE RECHERCHE

Définition : « **Enregistrements factuels (chiffres, textes, images, sons) utilisés comme source principale pour la recherche scientifique et généralement reconnus par la communauté scientifique comme nécessaires pour valider les résultats de la recherche.**

**Un ensemble de données de recherche constitue une représentation systématique et partielle du sujet faisant l'objet de la recherche. »** (OCDE, 2007, *[Principes et lignes directrices pour l'accès aux données de la recherche financée sur fonds publics](#)*)

En droit interne, les données de recherche trouvent leur protection principalement dans le

- **Code des Relations entre le Public et l'Administration,**
- **la loi pour une République Numérique du 7 octobre 2016,**
- **le Code de la Propriété Intellectuelle** et enfin
- **la loi Informatiques et Libertés du 6 janvier 1978.**

**MAIS**

**Il n'existe pas de droit unifié de la donnée.**

# LORSQUE LES DONNÉES SONT PERSONNELLES

Lorsque les données permettent de donner une information sur une personne physique alors leur utilisation est régie par le Règlement général de protection des données (RGPD) et par la loi Informatiques et Libertés du 6 janvier 1978 en France.

## ARTICLE 5 : Principes relatifs au traitement de données à caractère personnel

Les données à caractère personnel doivent être :

- traitées de manière **licite, loyale et transparente** au regard de la personne concernée ;
- collectées pour des **finalités déterminées, explicites et légitimes**, et **ne pas être traitées ultérieurement d'une manière incompatible avec ces finalités** (limitation des finalités);
- **adéquates, pertinentes et limitées à ce qui est nécessaire** au regard des finalités pour lesquelles elles sont traitées (minimisation des données) ;
- **exactes** et, si nécessaire, tenues à jour (principe d'exactitude)
- **Conservées pendant une durée limitée**
- Traitées de façon à garantir une **sécurité appropriée**

# L'OPEN DATA

---

La loi pour une République numérique a instauré plusieurs principes importants pour les données publiques :

- **Une ouverture par défaut** des données publiques et d'intérêt général
- Un accès sécurisé aux données **pour les chercheurs et les statisticiens**
- **Un libre accès** aux résultats des travaux de recherche publique et autorisation de la fouille de textes et de données

# MAIS LES DONNÉES SONT PROTÉGÉES PAR LA PI

Les données de recherche peuvent faire l'objet d'une protection supplémentaire : la propriété intellectuelle. Ainsi,

- les résultats de recherches peuvent faire être protégées par un brevet, si ces derniers représentent une invention et qu'un brevet a été conféré. (**L. 611-1 et s. CPI**)
- les données brutes ou informations peuvent faire l'objet d'une protection du droit *sui generis* des producteurs de données, branche du droit d'auteur (**L. 341-1 et s. CPI**)
- Les photos, vidéos, logiciels, écrits peuvent faire l'objet d'une protection par le droit d'auteur si ces derniers témoignent d'une originalité (**L. 111-2 CPI**).



**Le droit d'auteur ne protège pas les idées et les concepts et les algorithmes, ce sont leur formalisation qui est protégée (texte, logiciels,...)**

# LES DONNÉES DE RECHERCHE PUBLIQUES

## Les données sont considérées comme des documents administratifs.

- L'article **L. 300-2 du CRPA** définit les documents administratifs comme étant tous les documents « quels que soient leur forme et leur support ». Il donne une liste qui commence par “notamment” : elle n’est pas exhaustive et permet donc d’englober les données.
- De ce fait, comme les données sont donc des documents administratifs, cela implique un droit d’accès sur demande selon la procédure CADA (**Art. L. 311-1 CRPA**), un droit de diffusion (**Art. L. 312-1 CRPA**) et une obligation de diffusion (**Art. L. 312-1-1 CRPA**).
- Les données, en tant que documents administratifs, sont également considérées comme des informations publiques (**Art. L321-1 CRPA**) ce qui implique une libre réutilisation (même commerciale), sous conditions (**Art. L322-1 CRPA**) et gratuite (**Art. L324-4-1 CRPA**).

# LES DONNÉES DE RECHERCHE PUBLIQUES

## La communication des données publiques :

**Principe** : La loi n° 78-753 du 17 juillet 1978 (CADA), codifiée au livre III du code des relations entre le public et l'administration (CRPA), a institué **le principe de la liberté d'accès aux documents administratifs**. Ainsi, toute personne qui en fait la demande a le droit d'obtenir la communication de documents administratifs.

Depuis la **Loi pour une République Numérique**, les données achevées doivent obligatoirement être publiées en ligne si au moins l'un des **4 critères** suivants est rempli (**Art. L312-1-1 CRPA**) :

- Elles ont fait l'objet d'une demande de communication selon la procédure CADA.
- Elles figurent dans le répertoire des principaux documents mentionné au premier alinéa de l'article L. 322-6.
- Il s'agit d'une base de données.
- Les données ont un intérêt environnemental, social, sanitaire ou économique.

## 2 Exceptions du CRPA :

- Secret défense (**L. 311-5 CRPA**)
- Vie privée, secret médical ou secret d'affaires (**L. 311-6 CRPA**)

## Achèvement de la donnée :

**Principe** : Pas d'obligation de diffusion si les données ne sont pas achevées (**Art. L. 311-2 CRPA**)

**Exceptions** : 1) les données géographiques numérique qui relèvent de la directive européenne INSPIRE et codifiées aux articles **L. 127-1 à L. 127-10 du code de l'environnement.**

2) Les données personnelles qui doivent être communiquées sur demande de la personne concernée (**Art 15 RGPD**).

## Focus : les données géographiques numérique

Les autorités publiques créent et mettent à jour des métadonnées pour les séries et les services de données géographiques définis à l'article [L. 127-1](#) en conformité avec les modalités d'application définies dans le règlement (CE) n° 1205/2008 du 3 décembre 2008.

Ces métadonnées comprennent des informations relatives :

- a) A la conformité des séries de données géographiques avec les modalités d'application de l'interopérabilité mentionnées à la section 3 du présent chapitre ;
- b) Aux conditions applicables à l'accès et à l'utilisation des séries et des services de données géographiques et, le cas échéant, aux frais correspondants ;
- c) A la qualité et à la validité des séries de données géographiques ;
- d) Aux autorités publiques responsables de l'établissement, de la gestion, de la maintenance et de la diffusion des séries et des services de données géographiques ;
- e) Aux restrictions à l'accès public et aux raisons de ces restrictions.

# LES DONNÉES DE RECHERCHE PUBLIQUES

Depuis le 24 septembre 2023 :

**Le Règlement (UE) 2022/868 du 30 mai 2022 sur la gouvernance des données (DGA) -**

**Objectifs :**

- Facilite la réutilisation de données du secteur public
  - Crée les services d'intermédiation de données
- Crée les organismes altruistes en matière de données

→ Met en place un marché économique d'intermédiation de la donnée fondé sur des **tiers de confiance** n'utilisant pas eux-mêmes la donnée.

→ Le DGA définit **trois types de services d'intermédiation de données** à son article 10, dont l'objectif est d'établir une relation commerciale entre plusieurs types de personnes :

- entre détenteurs et utilisateurs de données (ex : plateformes d'échanges « B to B ») ;
- entre personnes concernées et les utilisateurs (ex : système de gestion des informations personnelles)
- les coopératives de données (ex : mutualisation des données en vue d'une gestion commune) ;

# LES DONNÉES DE RECHERCHE PUBLIQUES

**La mise à disposition des données** relève des organismes du secteur public (l'Etat, autorités locales ou régionales, organismes de droit public, associations formées par une ou plusieurs de ces autorités ou plusieurs de ces organismes).

- Pour être un organisme de droit public, il faut remplir **3 conditions** :
  - Avoir été créés pour satisfaire spécifiquement des besoins d'intérêt général et n'ont pas de caractère industriel ou commercial.
  - Être dotés de la personnalité publique.
  - Être financés majoritairement par un organisme du secteur public.

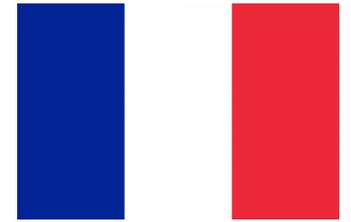
# LES DONNÉES DE RECHERCHE PUBLIQUES

Concernant la réutilisation de données, ce sont les organismes du secteur public qui autorisent ou refusent l'accès aux fins de réutilisation d'une ou plusieurs catégories de données.

Cette réutilisation doit respecter le droit européen et le droit interne (**Art 5 DGA**) comme respecter le RGPD, les droits PI, les secrets d'affaires...

Mais le droit sui generis des producteurs de base de données ne peut pas être opposées.

**DONC les entreprises privées peuvent réutiliser les données d'établissement public si elles y sont autorisées.**



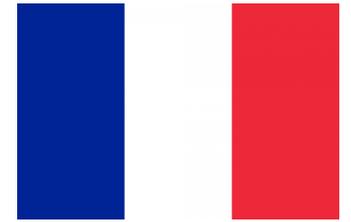
- Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés :

## **Art. 64 à 77**

Critère d'application territoriale spécifique (cumulatif avec les critères d'application territoriale RGPD)

S'applique lorsque les personnes concernées résident en France, même si le RT n'est pas établi en France

(article 5-1)



## ➤ Principes

### INTERDICTION

**ARTICLE 6 LIL :** Il est interdit de traiter des données à caractère personnel qui révèlent la prétendue origine raciale ou l'origine ethnique, les opinions politiques, les convictions religieuses ou philosophiques ou l'appartenance syndicale d'une personne physique ou de traiter **des données génétiques**, des données biométriques 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.

### INTERDICTION SAUF POURSUITE D'UNE FINALITE D'INTERET PUBLIC

**ARTICLE 66 LIL :** Les traitements de données à caractère personnel dans le domaine de la santé ne peuvent être mis en œuvre (hors exception) **qu'en considération de la finalité d'intérêt public qu'ils présentent.**

La garantie de normes élevées de qualité et de sécurité des soins de santé et des médicaments ou des dispositifs médicaux **constitue une finalité d'intérêt public.**

## ENCADREMENT DE L'ACTIVITE DE RECHERCHE

**ARTICLE 72 LIL :** Les traitements automatisés de données à caractère personnel **dont la finalité est ou devient la recherche ou les études dans le domaine de la santé ainsi que l'évaluation ou l'analyse des pratiques ou des activités de soins ou de prévention** sont soumis à la sous-section 1 de la présente section, sous réserve de la présente sous-section.

Le comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé peut se saisir ou être saisi, dans des conditions définies par décret en Conseil d'État, par la Commission nationale de l'informatique et des libertés ou le ministre chargé de la santé sur le caractère d'intérêt public que présentent les traitements mentionnés au premier alinéa du présent article.

**ARTICLE 73 LIL :** Des **méthodologies de référence** sont homologuées et publiées par la Commission nationale de l'informatique et des libertés.

Lorsque le traitement est conforme à une méthodologie de référence, il peut être mis en œuvre, **sans autorisation** mentionnée à l'article 66 de la présente loi, à la condition que son responsable adresse préalablement à la Commission nationale de l'informatique et des libertés une déclaration attestant de cette conformité.

### Droits de la personne

**ARTICLE 69 LIL:** Les personnes auprès desquelles sont recueillies des données à caractère personnel ou à propos desquelles de telles données sont transmises **sont individuellement informées** conformément aux dispositions du règlement (UE) 2016/679 du 27 avril 2016. (**voir article 71**)

**ARTICLE 74 LIL :** Toute personne a le droit de **s'opposer** à ce que des données à caractère personnel la concernant fassent l'objet de la levée du secret professionnel rendue nécessaire par un traitement de la nature de ceux mentionnés à l'article 65.

**Article 75 LIL:** Dans le cas où la recherche nécessite l'examen des caractéristiques génétiques, le **consentement éclairé et exprès des personnes concernées** doit être obtenu préalablement à la mise en œuvre du traitement de données. Le présent article n'est pas applicable aux recherches réalisées en application de l'article L. 1130-5 du code de la santé publique.

## CADRE LEGAL DE LA RECHERCHE TRAITANT DES DONNEES DE SANTE

### ➤ Recherche interne

□ **Comment:** à partir de données recueillies dans le cadre du suivi (thérapeutique ou médical) individuel des patients / et par les personnels assurant ce suivi ;/ et pour leur usage exclusif.

□ **Formalités:** Pour ce type de recherches, qui ne sont pas soumises au chapitre IX de la loi Informatique et Libertés, aucune formalité n'est requise si ce n'est: l'inscription au registre + information des personnes concernées.

### ➤ Recherche multicentrique

□ **Hypothèse n°1:** Si la recherche est strictement conforme à une méthodologie de référence (MR-001, MR-002, MR-003, MR-004, MR-005 ou MR-006 selon la typologie de la recherche), seul un engagement de conformité auprès de la CNIL est nécessaire.

□ **Hypothèse n°2:** Si la recherche n'est pas conforme à une méthodologie de référence, le responsable de traitement doit déposer une demande d'autorisation « recherche » (articles 72 et suivants de la loi Informatique et Libertés).

# DISPOSITIONS SPECIFIQUES TRAITEMENTS DONNEES DE SANTE

## CADRE LEGAL DE LA RECHERCHE TRAITANT DES DONNEES DE SANTE

### ➤ Recherche multicentrique / Hypothèse n°2

**2 cas doivent être distingués:** La recherche impliquant la recherche humaine (RIPH) / La recherche n'impliquant pas la recherche humaine (RNIPH)

#### Pour les recherches impliquant la personne humaine (RIPH) :

Un avis favorable du **CPP** (Comité de protection des personnes mentionné à l'article L. 1123-6 du code de la santé publique) est requis avant le début de la

recherche que ce soit dans le cas d'une conformité à une méthodologie de référence ou pour une autorisation « recherche ».

Dans ce dernier cas, l'avis favorable du CPP doit être obtenu avant de procéder à la demande d'autorisation et la recherche doit être conforme à une MR.

#### Pour les recherches n'impliquant pas la personne humaine (RNIPH) :

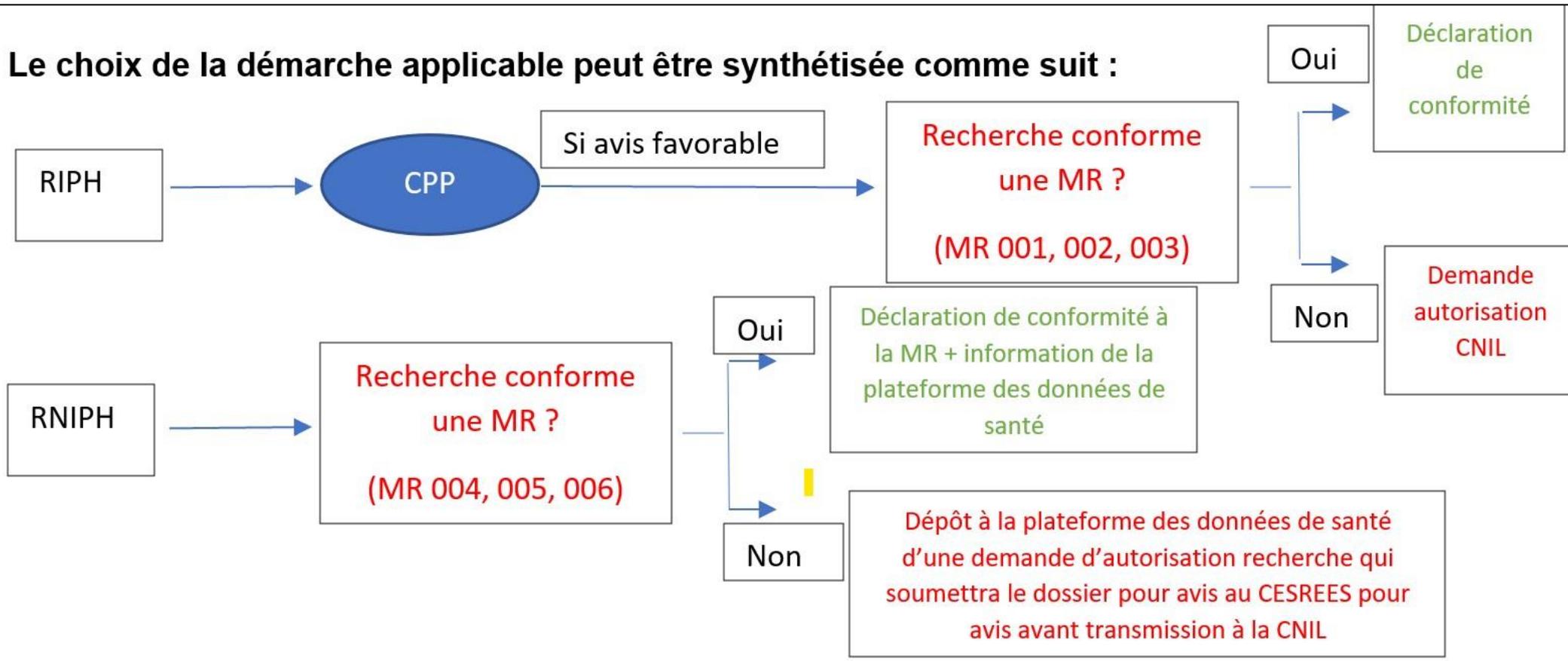
- le responsable de traitement doit déposer son dossier à l'INDS (Institut national des données de santé), qui se charge de le transmettre au **CESREES** (Comité d'Expertise pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé) pour avis, puis à la CNIL pour autorisation ;
- si la recherche est conforme à la MR-004, MR 005 ou MR 006, cette saisine de l'INDS et du CEREEES n'est pas requise mais des informations relatives à la recherche menée dans le cadre de ces MR devront être enregistrées dans un répertoire public, dont l'INDS est en charge.

# DISPOSITIONS SPECIFIQUES TRAITEMENTS DONNEES DE SANTE

## CADRE LEGAL DE LA RECHERCHE TRAITANT DES DONNEES DE SANTE

### ➤ Synthèse: Cadre de la recherche multicentrique

Le choix de la démarche applicable peut être synthétisée comme suit :



## CADRE LEGAL DE LA RECHERCHE TRAITANT DES DONNEES DE SANTE

### ➤ Liste des MR :

MR 001

• La méthodologie de référence MR-001 encadre les traitements comprenant des données de santé et présentant un caractère d'intérêt public, réalisés dans le cadre de recherches nécessitant le **recueil du consentement** de la personne concernée. Il s'agit plus précisément des recherches **interventionnelles**, y compris les recherches à risques et contraintes minimales, des essais cliniques de médicaments et des recherches nécessitant la réalisation d'un examen des caractéristiques génétiques. **L'information individuelle des patients est obligatoire**: Délibération n° 2018-153 du 3 mai 2018

MR 002

• La méthodologie de référence MR-002 concerne les **études non interventionnelles** de performances **ménées sur les dispositifs médicaux in vitro (DM DIV) en vue de leur mise sur le marché**.  
• Sont exclues du champ de cette méthodologie de référence les recherches biomédicales, les études faisant apparaître l'identité complète des personnes participant à l'étude, les études épidémiologiques, les études en génétique ayant pour objet d'identifier les personnes par leurs caractéristiques génétiques. Délibération n° 2015-256 du 16 juillet 2015

MR 003

• La méthodologie de référence MR-003 encadre les traitements comprenant des données de santé et présentant un caractère d'intérêt public, réalisés dans le cadre de **recherches impliquant la personne humaine** pour lesquelles la personne concernée ne s'oppose pas à participer après avoir été informée. Il s'agit plus précisément des **recherches non interventionnelles et des essais cliniques de médicaments par grappe**. L'information individuelle des patients est obligatoire. Délibération n° 2018-154 du 3 mai 2018

MR 004

• La méthodologie de référence MR-004 encadre les traitements de données à caractère personnel à **des fins d'étude, évaluation ou recherche n'impliquant pas la personne humaine**. La recherche doit présenter un **caractère d'intérêt public**. Délibération n° 2018-155 du 3 mai 2018

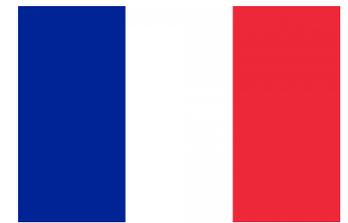
MR 005

• La méthodologie de référence MR-005 encadre l'accès par **des établissements de santé et des fédérations hospitalières** aux données du **Programme de médicalisation des systèmes d'information (PMSI) et aux RPU (Résumé de passage aux urgences)** mises à disposition sur la plateforme sécurisée de l'Agence technique de l'information sur l'hospitalisation (ATIH). Les études menées doivent présenter un caractère d'intérêt public et **aucun appariement avec d'autres données à caractère personnel n'est autorisé**. Les responsables de traitement doivent enregistrer leurs traitements auprès d'un répertoire public tenu par l'INDS. Délibération n° 2018-256 du 7 juin 2018

MR 006

• La MR 006 encadre l'accès par **des industriels de santé** aux données du Programme de médicalisation des systèmes d'information (PMSI) de l'Agence technique de l'information sur l'hospitalisation (ATIH) mises à disposition via une solution sécurisée. Les études menées doivent présenter un **caractère d'intérêt public** et aucun appariement avec d'autres données à caractère personnel n'est possible. Les responsables de traitement doivent enregistrer leurs traitements auprès d'un répertoire public tenu par l'INDS. Les industriels devront recourir à un bureau d'études/laboratoires de recherches ayant réalisé un engagement de conformité. Délibération n° 2018-257 du 7 juin 2018

# EXCEPTIONS



Section 3 de la LIL non applicable:

## Article 65 de la Loi Informatique et Libertés

Les traitements nécessaires aux fins **de la médecine préventive, des diagnostics médicaux, de l'administration de soins ou de traitements, ou de la gestion de services de santé** et mis en œuvre par un membre d'une profession de santé ou par une autre personne à laquelle s'impose en raison de ses fonctions l'obligation de secret professionnel.

Etudes conduites à partir des données recueillies lors du suivi médical ou individuel par les personnels assurant ce suivi et pour leur usage exclusif (« **étude interne** »)

Le **consentement** explicite

Traitements mis en œuvre pour l'exercice de leurs missions par les organismes chargés de la gestion d'un régime de base d'assurance maladie ainsi que la prise en charge des prestations par les organismes d'assurance maladie complémentaire

Sauvegarde des intérêts vitaux de la personne  
Les traitements mis en œuvre par **une association ou autre organisme à but non lucratif** si :

- le traitement se rapporte exclusivement aux membres de l'organisme ou personnes entretenant des contacts réguliers et
- la personne concernée a donné son consentement pour les données transmises hors de l'organisme

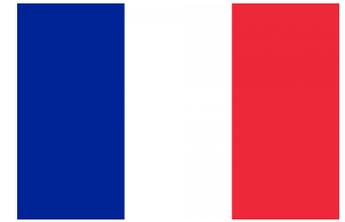
Données manifestement **rendues publiques** par la personne concernée  
Constatation, exercice ou défense **d'un droit en justice**

**Médecins responsables de l'information médicale (DIM)** dans les établissements de santé pour leurs missions

Traitements effectués par les **ARS et par l'Etat sur l'activité des établissements de santé**

Traitements mis en œuvre par **l'Etat** aux fins de **conception, de suivi ou d'évaluation des politiques publiques dans le domaine de la santé** ainsi que ceux réalisés aux fins de collecte, d'exploitation et de diffusion des statistiques dans ce domaine.

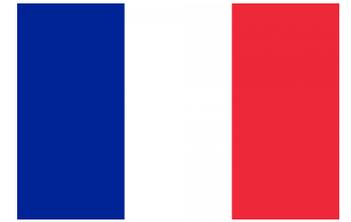
# QUELS SONT LES AUTRES RÉFÉRENTIELS À RESPECTER ?



## EN PLUS

- Dispositions sur le secret : **art. L. 1110-4 CSP**
- Dispositions relatives aux référentiels de sécurité et d'interopérabilité des données de santé : **art. L. 1110-4-1 CSP (PGSSI-S et CI-SIS)**
- Dispositions sur l'hébergement des données de santé : **art. L. 1111-8 et R. 1111-8-8 et s. CSP**
- Dispositions sur la mise à disposition des données de santé : **art. L. 1460-1 et s. du CSP**
- Interdiction de procéder à une cession ou à une exploitation commerciale des données de santé : **art. L. 1111-8 CSP, L. 4113-7 CSP**

# LES DONNÉES DE SANTÉ EN DROIT INTERNE



## ➤ Synthèse

Application du RGPD  
à tout traitement de  
données de santé

Si traitement soumis à  
exception de l'article 65,  
seul le RGPD s'applique

Si traitement non-conforme à une  
méthodologie de référence, obligation  
de demander l'autorisation de CNIL

Si traitement conforme à une  
MR, déclaration à la CNIL et à  
toutes autres autorités

+ Code de la santé publique

# LES ENTREPOTS DE DONNÉES DE SANTÉ

**CADRE :** Les entrepôts de données de santé sont des bases de données destinées à être utilisées à des fins de recherches, d'études ou d'évaluations dans le domaine de la santé.

Ces traitements sont soumis à l'autorisation de la CNIL sauf conformité au référentiel EDS.

## **RÉFÉRENTIEL EDS CNIL LIMITÉ CAR :**

- Mission d'intérêt public ou autorité publique dont est investi le RT
- Finalités déterminées (production d'indicateurs, amélioration dans le cadre du PMSI, fonctionnement d'aide au diagnostic médical ou à la prise en charge, réalisation d'étude de faisabilité)

**OBJECTIFS EDS :** Utilisation à des fins de recherche. Obligation de respecter les exigences pour chacun des projets de recherche envisagés (MR, autorisation Cnil)

**EXCLUSION:** La réutilisation des données de l'entrepôt dans le cadre de recherches, d'études ou d'évaluations dans le domaine de la santé doit faire l'objet des formalités à part (exemples : engagement de conformité aux méthodologies de référence ou demande d'autorisation recherche).

# LES ÉTABLISSEMENTS DE DONNÉES DE SANTÉ

**DESCRIPTION** : Le référentiel entrepôt ne s'applique qu'aux entrepôts de données de santé reposant sur l'exercice d'une mission d'intérêt public, au sens de l'article 6.1.e du RGPD.

L'entrepôt doit être nécessaire à l'exécution d'une mission d'intérêt public ou relevant de l'exercice de l'autorité publique dont est investi l'organisme responsable du traitement.

Un établissement privé de soins mettant en œuvre un entrepôt de données de santé nécessaire à l'exécution d'une mission d'intérêt public pourra être couvert par le référentiel.

Ce n'est pas le cas des entrepôts mis en œuvre par des entreprises privées à des fins de recherche sans exécution d'une mission d'intérêt public, sur le fondement de l'article 6.1.f du RGPD, par exemple.

Par ailleurs, certains entrepôts, tels que ceux reposant sur le recueil du consentement des personnes concernées ou ceux qui ont exclusivement pour objectif de dématérialiser les dossiers médicaux ne sont pas soumis à autorisation (article 65 de la loi Informatique et Libertés).

Concernant les entrepôts de données de santé **qui ne répondent pas aux critères du référentiel**, les organismes doivent solliciter la CNIL afin d'obtenir **une autorisation individuelle** préalablement à la mise en œuvre de leur entrepôt.

# DÉFINITION

---

**Anonymisation** : Traitement de données personnelles permettant de rendre impossible l'identification d'une personne à partir d'un jeu de données et permet, ainsi, de respecter sa vie privée

**Une donnée anonyme n'est plus une donnée personnelle**  
**Donnée anonyme = donnée non soumise au RGPD**

**Pseudonymisation** : Traitement de données personnelles réalisé de manière à ce qu'on ne puisse plus attribuer les données à une personne physique identifiée sans information supplémentaire

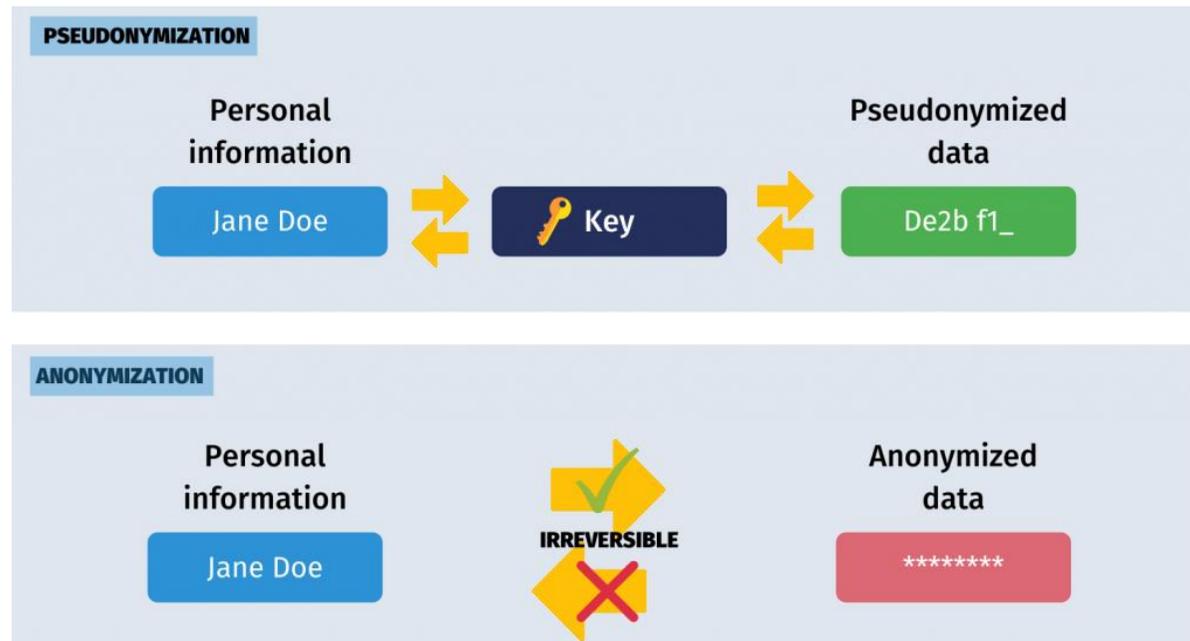
**Une donnée pseudonyme est une donnée personnelle**  
**Donnée pseudonyme = donnée soumise au RGPD**

**L'anonymisation est donc un enjeu crucial permettant de perdre la qualification de donnée personnelle et de se soustraire à toute la réglementation applicable !**

# DÉFINITION

## Anonymisation $\neq$ Pseudonymisation

- Une anonymisation implique une opération irréversible
- Une pseudonymisation permet la ré-identification des personnes



# INTÉRÊT DE L'ANONYMISATION

L'anonymisation ouvre des potentiels de **réutilisation des données initialement interdits** du fait du caractère personnel des données exploitées, et permet ainsi aux acteurs d'exploiter et de partager leur « gisement » de données **sans porter atteinte à la vie privée des personnes**. En devenant une donnée brute, la donnée n'est plus règlementée par les dispositions du RGPD

*Source CNIL, 19 mai 2020.*

**L'anonymisation des données, en vue de leur réutilisation, devient un enjeu majeur :**

## **La start-up Octopize, spécialisée dans l'anonymisation des données, séduit le CHU de Brest**

La start-up rennaise Octopize a développé une technologie d'anonymisation des données de santé pour faciliter leur réutilisation. Elle a été sélectionnée par le CHU de Brest et Roche pour simplifier la transmission des données de cohortes des patients dans le cadre d'études cliniques.

[L'usine digitale](#)

# FLOU SUR LES CRITÈRES DE DISTINCTIONS

---

## **Arrêt CJUE c/ Deloitte, 26 avril 2023 :**

Hypothèse dans laquelle une table de correspondance existe → a priori, ici il ne s'agit pas de données anonymes dans la mesure où il est possible de ré-identifier les personnes concernées.

Toutefois, dans le cas d'espèce, la personne qui dispose des données pseudonymisées n'a pas la table de correspondance en sa possession | **La CJUE considère les données comme anonymes puisque la personne n'a pas la clé pour déchiffrer ces données**



**Arrêt remarqué et très commenté, car ouvrant potentiellement une portée plus importante à la définition de données anonymes.**

**Attention toutefois :**

- **Arrêt justifié par les circonstances particulières de l'espèce**
- **Arrêt qui ne concerne pas des données de santé**

# TRAITEMENT D'ANONYMISATION

---

## RGPD

### Considérant 26 RGPD:

*« Pour déterminer si une personne physique est identifiable, il convient de prendre en considération l'ensemble des moyens raisonnablement susceptibles d'être utilisés par le responsable du traitement ou par toute autre personne pour identifier la personne physique directement ou indirectement, tels que le ciblage »*

### Sources:

- [Avis 05/2014 de l'article 29 sur les Technique d'anonymisation](#)
- [Ressources CNIL: L'anonymisation des données personnelles](#)

# TRAITEMENT D'ANONYMISATION

---

**Définition:** 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. Les données anonymisées ne sont plus considérées comme personnelles.

## **Différence avec la pseudonymisation:**

La pseudonymisation est un traitement de données personnelles réalisé de manière à empêcher toute identification directe, typiquement en remplaçant les informations directement identifiantes par des numéros non significatifs.

À l'inverse de l'anonymisation, la pseudonymisation ne permet donc pas d'empêcher l'identification indirecte, et reste généralement réversible : les données restent personnelles.

**Choix:** Le choix entre anonymisation et pseudonymisation résulte d'un compromis entre les questions suivantes :

1. Est-il nécessaire de préserver la finesse des données au niveau des individus composant le jeu de données ?
2. Est-il nécessaire de pouvoir partager simplement et largement le jeu de données ?

# TRAITEMENT D'ANONYMISATION

**Quand choisir l'anonymisation ?** Quand l'utilisation ne justifie pas le recours à des données identifiantes ou qu'il est plus important de permettre un partage large et simple des données (car n'ayant plus à respecter la législation relative à la protection des données personnelles).

Exemples d'utilisation :

- Statistiques génériques (e.g. populationnelles)
- Exploitation large via des data challenges
- Partage libre de données pour assurer la reproductibilité de la recherche par exemple Choisir

Exemple de jeu de données anonymisé:

Drug ID	Age group	Period	Patient count
A	20-30	Q1 2023	225
A	30-40	Q1 2023	137
B	20-30	Q2 2023	51

# TRAITEMENT D'ANONYMISATION

**Quand choisir la pseudonymisation?** Quand il est plus important de préserver la finesse des informations individuelles composant le jeu de données tout en limitant les risques liés à leur traitement.

Exemples d'utilisation :

- Études nécessitant des données fines à l'échelle individuelle (e.g. étude de parcours de soin individuels)
- Appariement de jeux de données à l'aide d'informations individuelles (e.g. NIR masqués)

Exemple de jeu de données pseudonomisé:

Patient ID	Birth Date	Drug ID	Date
EJ21004D	1993/07	A	2020/01/04
EJ21004D	1993/07	B	2020/05/23
FE040JO	1974/01	A	20202/01/06

# TRAITEMENT D'ANONYMISATION

---

**Comment anonymiser?** Le choix du processus d'anonymisation dépend des réutilisations futures envisagées, et peut donc considérablement varier d'un projet à l'autre et d'un jeu de données à un autre. On peut distinguer deux familles de techniques fréquemment utilisées pour anonymiser des données : la généralisation et la randomisation.

**La généralisation** altère la finesse des données en modifiant leur échelle: k-anonymat l-diversité t-proximité

**La randomisation** altère la véracité des données pour affaiblir le lien avec l'individu : Ajout de bruit, Confidentialité différentielle, Permutation.

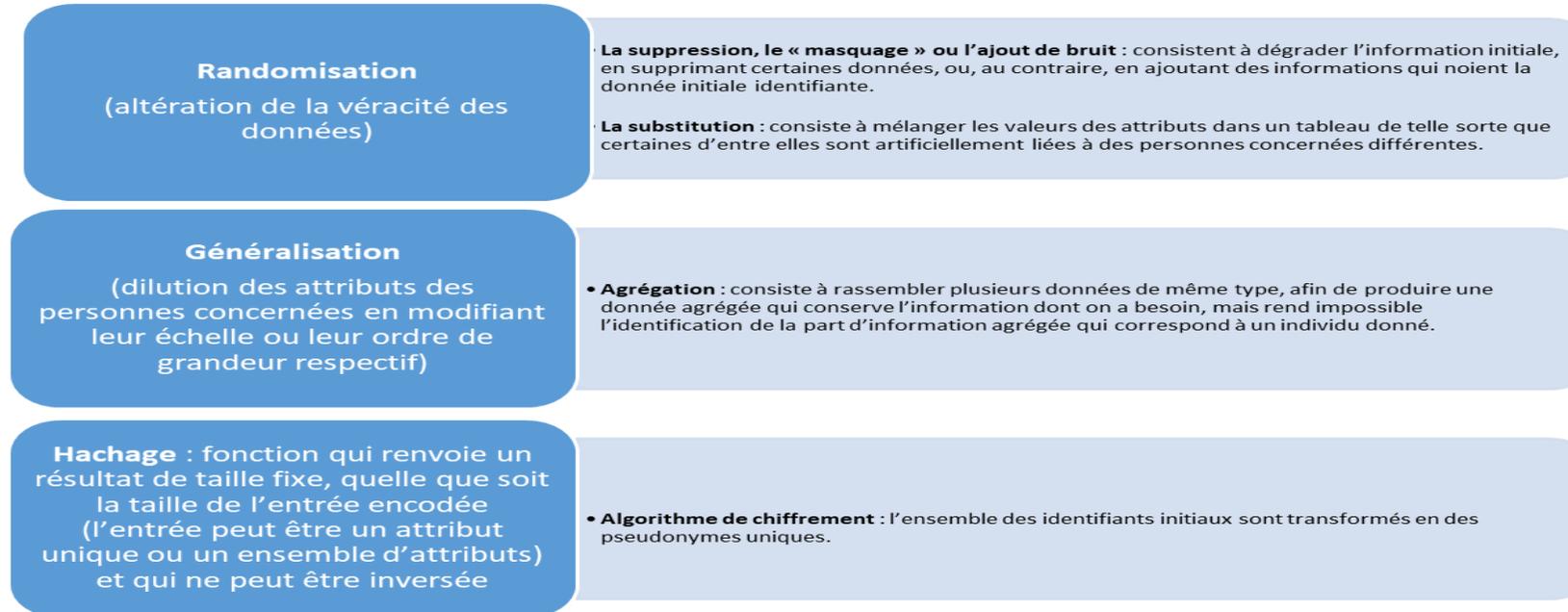
**Évaluation des techniques d'anonymisation.** Il ressort de l'avis du G29 qu'aucune technique d'anonymisation n'est en théorie infaillible.

Toutefois, pour aider à évaluer une bonne solution d'anonymisation, le G29 propose trois critères :

- L'individualisation : est-il toujours possible d'isoler un individu ?
- La corrélation : est-il possible de relier entre eux des ensembles de données distincts concernant un même individu ?
- L'inférence : peut-on déduire de l'information sur un individu ?

# TRAITEMENT D'ANONYMISATION

## Comment anonymiser?



- Parallèlement à la recherche de nouvelles techniques d'anonymisation, des recherches se tournent vers d'autres voies, telles que **l'exploration des techniques homomorphes ou la conception d'algorithmes de génération de données de synthèse anonymes** – avatars – à partir de données personnelles.
- **Mais il n'existe pas encore actuellement de procédure d'anonymisation certifiée, en particulier par la CNIL.**

# TRAITEMENT D'ANONYMISATION

---

**Attention, la CNIL et le G29 précise bien que le procédé d'anonymisation constitue en soi un traitement ultérieur de données à caractère personnel :**

*« L'anonymisation est le **résultat du traitement des données personnelles** afin d'empêcher, de façon irréversible, toute identification »*

*« **L'anonymisation constitue un traitement ultérieur des données à caractère personnel**; à ce titre, elle doit satisfaire à l'exigence de compatibilité au regard des motifs juridiques et des circonstances du traitement ultérieur ».*

[Avis du G29 sur les techniques d'anonymisation](#)

## **Conséquences :**

**Le traitement d'anonymisation est un traitement de données à caractère personnel et les principes relatifs au RGPD doivent être respectés :**

- Mesure de sécurité suffisantes
- Réalisation d'une analyse d'impact
- ...

# L'EUROPEAN HEALTH DATA SPACE

---

L'Espace européen des données de santé est **une proposition réglementaire** à l'initiative de la Commission européenne **le 3 mai 2022** qui:

- vise à offrir aux citoyens de l'UE un meilleur contrôle sur leurs données de santé personnelles
- soutient l'utilisation des données de santé pour améliorer la prestation des soins de santé, la recherche, l'innovation et l'élaboration des politiques
- permet à l'UE de tirer pleinement parti du potentiel offert par un échange, une utilisation et une réutilisation sûrs et sécurisés des données de santé.

# L'EUROPEAN HEALTH DATA SPACE

---

## A qui s'applique l'European Health Data Space ?

«**détenteur de données**», toute personne physique ou morale qui est une entité ou un organisme du secteur de la santé ou des soins ou qui effectue des recherches dans ces secteurs, ainsi que les institutions, organes et organismes de l'Union qui ont le droit ou l'obligation, conformément au présent règlement, au droit de l'Union applicable ou à la législation nationale le mettant en œuvre, ou, dans le cas de données à caractère non personnel, par le contrôle de la conception technique d'un produit et de services liés, la capacité de mettre à disposition certaines données, y compris de les enregistrer, de les fournir, d'en restreindre l'accès ou de les échanger;

→ **Contrairement au DGA, l'espace européen de données de santé ne restreint pas aux organismes publics.**

# L'EUROPEAN HEALTH DATA SPACE

---

## Objectifs sur la réutilisation secondaire des données de santé :

Mettre en place un dispositif cohérent, fiable et efficace pour l'utilisation des données de santé à des fins de recherche, d'innovation, d'élaboration des politiques et de réglementation.

**Dans des conditions strictes, les chercheurs, les innovateurs, les institutions publiques ou les entreprises auront accès à de grandes quantités de données de santé de qualité élevée.**

## Modalités d'accès aux données :

Pour accéder à ces données, les chercheurs, les entreprises ou les institutions devront demander une **autorisation auprès de l'organisme responsable de l'accès aux données de santé**, qui devra être mis en place dans chaque État membre.

L'accès ne sera autorisé que si les données demandées sont utilisées à des **fins particulières, dans des environnements fermés et sécurisés et sans que l'identité des personnes ne soit révélée.**

# L'EUROPEAN HEALTH DATA SPACE

---

**ART 33 :** établit plusieurs catégories de données que leur détenteur met à disposition à des fins d'utilisation secondaire. Il s'agit notamment des données ayant une incidence sur la santé, des données génomiques sur les pathogènes pertinentes ayant une incidence sur la santé humaine, des données administratives relatives à la santé ou encore données de santé électronique générées par la personne.

**ART 34 :** Finalités pour lesquelles les données de santé peuvent être traitées

- Intérêt public
- Aider les organismes du secteur public à accomplir les tâches
- Produire des statistiques officielles
- Activités d'éducation ou d'enseignement
- Activité de développement et d'innovation pour des activités contribuant à la santé publique
- La formation, l'expérimentation d'algorithmes contribuant à la santé publique
- Fournir des soins de santé personnalisés

# L'EUROPEAN HEALTH DATA SPACE

---

## La réutilisation des données de santé section 3 :

**L'article 44** impose le respect du principe de minimisation des données dans le cadre de leur réutilisation pour les finalités prévues par le texte. Il prévoit notamment que :

- L'organisme responsable de l'accès aux données de santé veille à ce que l'accès soit accordé uniquement aux données de santé électroniques demandées pertinentes pour la finalité.
- Les organismes responsables de l'accès aux données de santé donnent accès aux données de santé électroniques dans un format anonymisé.

**L'article 45** prévoit que toute personne physique ou morale peut présenter une demande d'accès aux données et **l'article 46** dispose que les organismes responsables de l'accès aux données de santé évaluent la demande.

# INTERVENANTE

---



**ELISE DUFOUR**  
**AVOCAT ASSOCIE**  
edufour@bignonlebray.com



MERCI



PARIS

75, rue de Tocqueville  
75017 Paris

