CNIL.
COMMISSION NATIONALE
INFORMATIQUE & LIBERTÉS

PROTÉGER les données personnelles
ACCOMPAGNER l'innovation
PRÉSERVER les libertés individuelles

Essais in silico et traitements de données personnelles

Journée du consortium 6 « Santé in silico »

Manon de FALLOIS, adjointe à la cheffe du service de la santé

9 février2024

Plan de la présentation

1. Présentation du droit applicable à la protection des données
2. Les principales notions en matière de traitements de données et les principes clés
3. Les formalités applicables à certains traitements de données de santé
4. Recherches en santé : les bons réflexes à avoir
5. Le cycle de vie d'une IA en santé: l'exemple des essais in silico

Le droit applicable à la protection des données

- **Règlement général sur la protection des données** (entré en vigueur le 25 mai 2018)
- **Loi n°78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés** dite « *informatique et libertés* » modifiée
- **Décret n°2019-536 du 29 mai 2019 modifié pris pour l'application de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés**
- **En cas de traitement de données du SNDS :**
  - **Loi n°2016-41 du 26 janvier 2016 de modernisation de notre système de santé**
  - **Loi n° 2019-774 du 24 juillet 2019 relative à l'organisation et à la transformation du système de santé**
  - **Décret n° 2021-848 du 29 juin 2021 relatif au traitement de données à caractère personnel dénommé « système national des données de santé »**
- **Autres dispositions légales (code pénal, code de la santé publique, code civil…)**

+ Règlement européen IA act (adopté le 2 février 2024, entrée en vigueur en 2025)

Les principales notions

Manon de Fallois, adjointe à la cheffe du service de la santé (CNIL)

La notion de traitement de données à caractère personnel

Données à caractère personnel

Toute information se rapportant à une **personne physique identifiée** ou **identifiable** directement ou indirectement.

- Directement identifiant
- Indirectement identifiant
- Recoupement d'informations

Traitement

Toute **opération** portant sur des données personnelles, **quel que soit le procédé utilisé.**

Par exemple:

- enregistrer,
- organiser,
- conserver,
- modifier,
- transmettre,
- etc.

Données pseudonymisées et données anonymisées: quelle distinction?

Données pseudonymisées

Pseudonymisation: *« le traitement de données à caractère personnel de telle façon que celles-ci ne puissent plus être attribuées à une personne concernée précise sans avoir recours à des informations supplémentaires, pour autant que ces informations supplémentaires soient conservées séparément et soumises à des mesures techniques et organisationnelles afin de garantir que les données à caractère personnel ne sont pas attribuées à une personne physique identifiée ou identifiable »* (article 4 du RGPD)

Une donnée non directement identifiante peut être une donnée à caractère personnel : **donnée pseudonymisée / codée (la plupart du temps, en recherche)**

Données anonymisées

Position du G29 (avis 05/2014 sur les techniques d'anonymisation)

« Une solution d'anonymisation doit être construite au cas par cas et adaptée aux usages prévus. 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 ?*

Ainsi :

*un **ensemble de données pour lequel il n'est possible ni d'individualiser ni de corréler ni d'inférer est a priori anonyme ;***

un ensemble de données pour lequel au moins un des trois critères n'est pas respecté ne pourra être considéré comme anonyme qu'à la suite d'une analyse détaillée des risques de ré-identification. »

Une donnée « anonyme » n'est PLUS/PAS une donnée à caractère personnel

Responsable de traitement et sous-traitant: quelle distinction?

Responsable de traitement

« *la* ***personne physique ou morale****, l'autorité publique, le service ou un autre organisme qui,* ***seul ou conjointement avec d'autres****, détermine les finalités et les moyens du traitement* » (article 4 du RGPD)

Exemples :

- ✓ Secteur privé : la société représentée par son président
- ✓ Secteur public : l'hôpital représenté par son directeur

En cas de **responsabilité conjointe de traitement** : nécessité pour les responsables conjoints de définir de manière transparente leurs obligations respectives (article 26 du RGPD).

Sous-traitant

« *la personne physique ou morale, l'autorité publique, le service ou un autre organisme qui traite des données à caractère personnel pour le compte du responsable du traitement* » (article 4 du RGPD)

En résumé: le sous-traitant agit sous l'autorité du responsable de traitement et sur ses instructions.

Conclusion d'un contrat ou d'un acte juridique avec le sous-traitant (article 28 du RGPD).

Pour en savoir plus : voir les lignes directrices européennes concernant les notions de responsable du traitement et de sous-traitant dans le RGPD.

Fournisseur et utilisateurs de l'IA: quelle distinction ?

Fournisseur

La personne physique ou morale, l'autorité publique, une agence ou un autre organisme qui développe ou fait développer un système d'IA en vue de le mettre sur le marché ou de le mettre en service sous son propre nom ou sa propre marque, que ce soit contre rémunération ou gratuitement.

Utilisateur

Toute personne physique ou morale, autorité publique, agence ou autre organisation utilisant un système d'IA sous sa propre autorité, sauf lorsque le système est utilisé dans le cadre d'une activité personnelle non commerciale.

Utilisateur final

La personne concernée par le système.

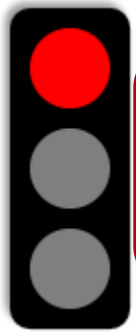
Pour en savoir plus : voir la fiche publiée sur le site web de la CNIL.

Qu'est-ce qu'une donnée de santé ?

Article 4 du RGPD *« données relatives à la **santé physique ou mentale**, passée, présente ou future, d'une personne physique (y compris la prestation de services de soins de santé) qui révèlent des informations sur l'état de santé de cette personne »*

Pour en savoir plus : voir la fiche publiée sur le site web de la CNIL.

Les données de santé, une catégorie particulière de données



<u>Interdiction</u> de traiter des données relatives à la santé
(article 9-I du RGPD et article 6 LIL)

Pour traiter des données de santé, il faut justifier de l'une des **exceptions** *(article 9-I du RGPD et article 6 LIL)*

Autres catégories de données sensibles: origine ethnique, vie sexuelle, orientation sexuelle, données génétiques, opinions politiques, convictions religieuses, etc.

Exemples:

- Le **consentement** explicite
- **Obligations** liées au droit du travail, protection sociale, sécurité sociale
- Sauvegarde des **intérêts vitaux** de la personne
- **Motifs d'intérêt public important**
- **Médecine préventive, diagnostics médicaux**, prise en charge sanitaire ou sociale ou **gestion des systèmes et services de soins en santé**
- **Motifs d'intérêt public** dans le domaine de la **santé publique**
- **Recherche scientifique**, fins archivistiques ou statistiques

Les principes clés en cas de traitement de données

01

Licéité, loyauté & transparence du traitement

02

Finalité déterminée, explicite et légitime

03

Minimisation des données

Exactitude des données collectées

04

05

Durée de conservation limitée

Pour en savoir plus : voir le guide pratique sur les durées de conservation.

Intégrité
Confidentialité
Disponibilité

06

Principaux risques : accès illégitime, modification non désirée, disparition

07

Respect des droits des personnes

Conformité

Les droits des personnes concernées

1. Droit à la transparence
2. Droit d'accès
3. Droit de rectification
4. Droit d'opposition
5. Droit à l'effacement
6. Droit à la limitation
7. Droit à la portabilité des données
8. Décision individuelle automatisée

Décision individuelle automatisée (1)

- Décision prise à l'égard d'une personne, par le biais d'algorithmes appliqués à ses données personnelles, sans qu'aucun être humain n'intervienne dans le processus.

- Nombreux domaines d'activités concernés (finance, fiscalité, marketing…)

- **Par principe, les personnes ont le droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé et produisant des effets juridiques les concernant (impact pour les droits et libertés) ou l'affectant de manière significative de façon similaire.**

Décision individuelle automatisée (2)

- **Par exception**, une personne peut faire l'objet d'une décision entièrement automatisée, même si cette dernière a un effet juridique ou un impact significatif sur elle dans certains cas particuliers (décisions fondées sur le consentement explicite des personnes, décisions nécessaires à la conclusion ou à l'exécution d'un contrat, décision prévue par le droit de l'Union européenne ou de l'Etat membre…).
- Dans cette hypothèse, les personnes disposent de **droits supplémentaires** lorsqu'une décision entièrement automatisée est prise à leur égard et les affecte particulièrement:
  - Obligations spécifiques de transparence (information sur l'existence de cette décision, de sa logique sous-jacente, de l'importance et des conséquences de cette décision);
  - Droit à une intervention humaine (pour réexaminer la situation, donner son point de vue, obtenir une explication sur la décision ou la contester).
- Restrictions encore plus fortes quand il s'agit de données sensibles (ex: données de santé).

Pour en savoir plus : voir la fiche dédiée disponible sur le site web de la CNIL.

Les formalités applicables à certains traitements de données de santé

Manon de Fallois, adjointe à la cheffe du service de la santé (CNIL)

Dans quels cas faut-il réaliser une formalité auprès de la CNIL? (1)

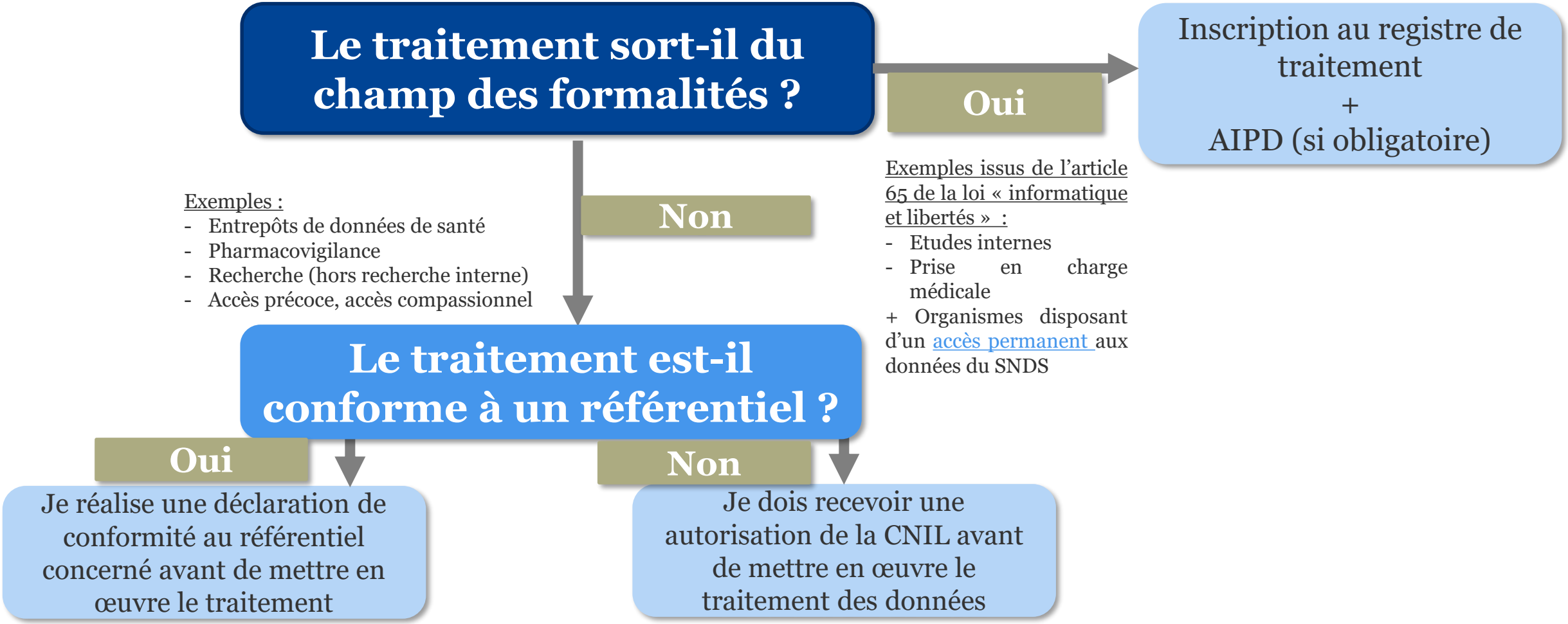
Traitements exempts de formalités – article 65 de la Loi « informatique et libertés »
(liste non exhaustive)

- **médecine préventive, diagnostics médicaux, administration des soins, mis en œuvre par un professionnel de santé**
- **consentement explicite**
- sauvegarde des **intérêts vitaux** (si incapacité physique ou juridique de donner son consentement)
- données manifestement **rendues publiques** par la personne concernée
- **études dites « internes »** à usage interne

Traitements soumis à formalités préalables

- Les traitements présentant une finalité d'intérêt public qui ne sont **pas exemptés de formalités**
- Les traitements automatisés 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

Dans quels cas faut-il réaliser une formalité auprès de la CNIL? (2)



Recherches en santé: les bons réflexes à avoir

Manon de Fallois, adjointe à la cheffe du service de la santé (CNIL)

Qu'est-ce qu'un entrepôt de données de santé et comment le distinguer d'une recherche?

La notion « d'entrepôt »

Les entrepôts de données de santé sont créés principalement **pour collecter et disposer de données massives**

- **Origine variée des données** (données relatives à la prise en charge médicale du patient, données socio-démographiques, données issues de précédentes recherches etc.)
- **Longue prolongée** de l'entrepôt
- **Base de données alimentée au fil de l'eau**
- Données réutilisées à la réalisation de **traitements ultérieurs**

Entrepôt	Recherche
Permet la réalisation ultérieure d'un nombre important de projets (dont les finalités sont diverses)	Finalité précise et répond à une question de recherche scientifique spécifique et ponctuelle
Constitué afin d'obtenir un volume de données important.	Les données sont collectées spécifiquement pour les besoins du projet.
Constitué pour une durée assez longue (10 ans en général)	La durée de la recherche est limitée et connue

Pour en savoir plus : fiche pratique ***« Traitements de données de santé : comment faire la distinction entre un entrepôt et une recherche et quelles conséquences ? »***

Existence de travaux en cours au sujet du statut des cohortes et sur la mise en conformité des registres

Présentation des étapes « clés »

Attention: recherche en santé = régime spécifique

1. Identifier la nature de la recherche

- ❑ Recherche impliquant la personne humaine
- ❑ Recherche n'impliquant pas la personne humaine

2. Identifier le périmètre de la recherche

- ❑ Recherche interne
- ❑ Recherche multicentrique

3. Procéder aux ajustements nécessaires

Avant la mise en œuvre de l'étude, il est nécessaire de corriger les points de non-conformité. Une attention particulière doit être apportée à l'information et à la sécurité des données.

4. Réaliser les démarches

Les démarches (autorisation CNIL, saisine comité éthique, etc.) varient en fonction de la nature (étape 1) et du périmètre (étape 2) de la recherche.

Pour aller plus loin

2 fiches pratiques dans la rubrique santé (cnil.fr)
- « *Recherche en santé : quel est le cadre légal ?* »
- « *Comment procéder pour une thèse ou un mémoire ?* »

La qualification de la recherche

Recherche impliquant la personne humaine (RIPH)

« *Recherches organisées et pratiquées sur des* ***personnes volontaires saines ou malades****, en vue du* ***développement des connaissances biologiques ou médicales*** *qui visent à évaluer :*

1° Les mécanismes de fonctionnement de l'organisme humain, normal ou pathologique ;
2° L'efficacité et la sécurité de la réalisation d'actes ou de l'utilisation ou de l'administration de produits dans un but de diagnostic, de traitement ou de prévention d'états pathologiques. (CSP, art. R.1121-1) »

Recherche n'impliquant pas la personne humaine (RNIPH)

Collecte de données supplémentaires pour les besoins de la recherche sans répondre à la définition de RIPH (notamment la finalité).

Réutilisation (changement de finalité) de données déjà acquises [par exemple, les données issues de bases médico-administratives (ex: SNDS) ou d'un registre agréé, d'entrepôt de données ou de dossiers médicaux sans que de nouvelles informations soient collectées auprès des personnes concernées pour les besoins de la recherche].

Le périmètre de la recherche: le cas de la recherche « interne »

Une recherche est considérée comme « **interne** » si elle est menée:

- 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.

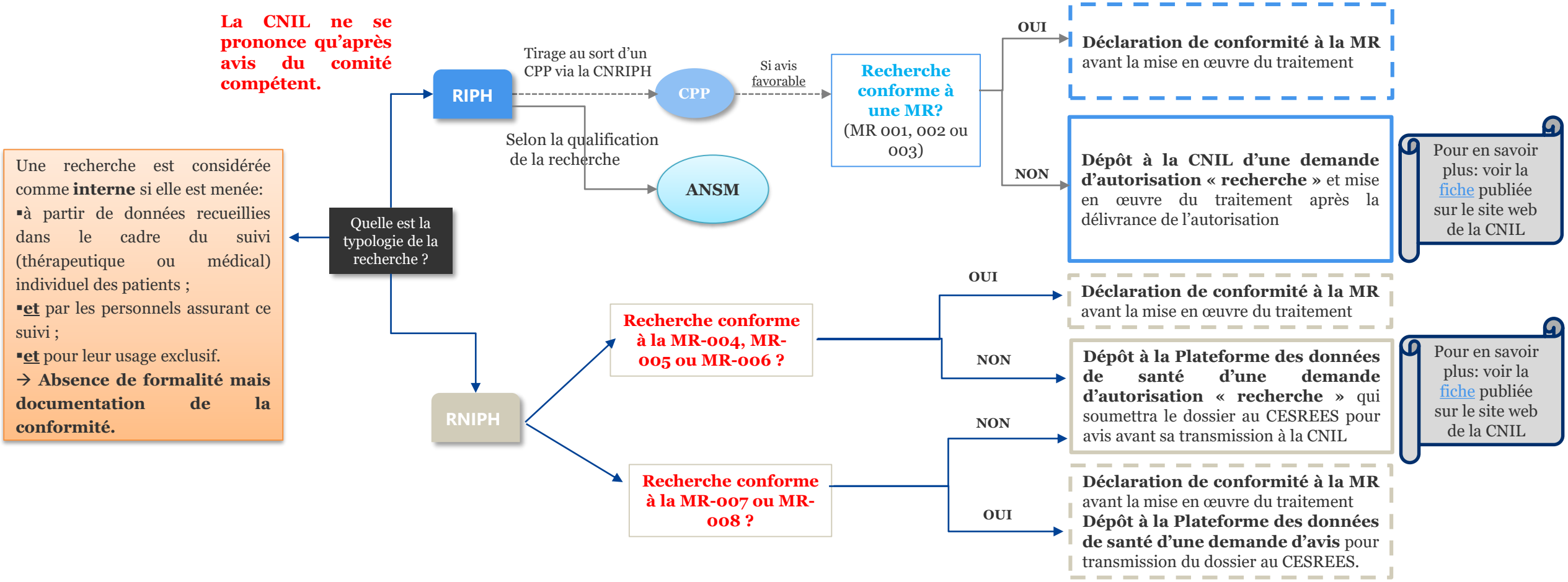
❖ **Exemple**

Une recherche menée par un interne agissant sous la responsabilité d'un chef de service sur les patients qu'il a suivis dans son service et uniquement sur les données collectées lors de leur prise en charge.

Pas de formalité vis-à-vis de la CNIL pour les études internes mais :

- Inscription au registre des activités de traitement
- Nécessité de respecter les droits des personnes concernées
- Nécessité de mettre en place des mesures de sécurité adéquates.

Recherche en santé réalisées en France : les démarches en synthèse



Cycle de vie d'une IA en santé: l'exemple des cohortes de patients virtuels

Manon de Fallois, adjointe à la cheffe du service de la santé (CNIL)

Formalités pour les traitements en santé par une IA

0. Cycle de vie d'une IA en santé

L'IA n'existe pas encore et est **en cours de développement**

L'IA est utilisée selon l'objectif pour lequel elle a été créée de manière opérationnelle **(phase de déploiement)**

Formalités pour les traitements en santé par une IA

1. Le développement de l'algorithme: la phase de développement

Assimilable à un traitement de recherche dans le domaine de la santé

Formalités applicables aux recherches dans le domaine de la santé

La déclaration de conformité à un référentiel

- **MR-001, MR-002, MR-003, MR-004, MR-005, MR-006,**
- **MR-007, MR-008 (avec un avis favorable du CESREES)**

OU

Le dépôt d'une demande d'autorisation (en l'absence de conformité à un référentiel)

Ma recherche implique la personne humaine :

après avis du comité de protection des personnes, dépôt d'une demande d'autorisation « recherche » auprès de la CNIL.

Ma recherche n'implique pas la personne humaine :

dépôt d'une demande d'autorisation « recherche » auprès de la Plateforme des données de santé.

1. Le développement de l'algorithme (1)

Cas d'usage : réutilisation de données de santé déjà collectées pour développer le système d'IA générant les cohortes de patients virtuels :

- qualification de **recherche n'impliquant pas la personne humaine** :
  - déclaration de conformité à la méthodologie de référence MR-004 ;
  - dépôt d'une demande d'autorisation après avis du Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé
- **respect des « principes clés »** du RGPD et de la loi « informatique et libertés », et notamment de la nécessité de poursuite d'une finalité d'intérêt public.

1. Le développement de l'algorithme (2)

- Identification :
  - du caractère anonyme ou non des données traitées (niveau de granularité des données)
  - du rôle des organismes (fournisseurs, utilisateurs de systèmes d'IA…) et leur qualification au regard du RGPD
  - de l'objectif (finalité du traitement)
  - de la proportionnalité des techniques choisies
  - de son impact et des interactions avec les personnes concernées
  - des avantages du recours au système d'IA

Pour en savoir plus: voir la fiche publiée sur le site web de la CNIL

1. Le développement de l'algorithme (3)

- En cas de réutilisation de données :
  - vérification que la base de données initiale a été constituée conformément à la règlementation
  - vérification que la base de données initiale est de qualité / représentative pour éviter les biais
- Documentation de la conformité (AIPD, analyse des risques de réidentification…) et notamment:
  - identification de la base légale du traitement et l'exception pour traiter les données sensibles
  - respect du principe de minimisation
  - modalités d'information et d'exercice des droits (information, non opposition, explicabilité…)
  - mise en place de mesures techniques et organisationnelles appropriées
- Documentation de la méthodologie (hypothèses effectuées sur les données d'entraînement, réalisation d'une étude des biais) pour faire en sorte d'avoir un traitement fiable et robuste (en lien avec l'intérêt public)

Pour en savoir plus: voir la fiche publiée sur le site web de la CNIL

1. Le développement de l'algorithme (4)

- conception et développement d'un algorithme fiable
  - identification du type d'algorithme,
  - justification de son choix et de ses modalités de fonctionnement (revue de la littérature, tests, comparaisons, suivi de la mise à jour, respect de l'état de l'art)
- vérifier la qualité du système en environnement contrôlé (intégration de tests de validation)

En construction :

- Quelles modalités d'évaluation *a priori* par les comités d'éthique et les régulateurs ?
- Quelles modalités d'évaluation *a posteriori* (notamment le principe de garantie humaine de l'IA) ?

Pour en savoir plus: voir la fiche publiée sur le site web de la CNIL

Formalités pour les traitements en santé par une IA

2. Utilisation de l'IA de manière opérationnelle

L'IA est incluse dans le traitement principal

Exception à formalités (article 65 LIL)

Utilisation d'un algorithme pour faire un diagnostic par le médecin qui prend en charge le patient :

- Le responsable de traitement est le professionnel de santé (ou l'établissement de santé) ;
- L'éditeur de logiciel agit en tant que sous-traitant.

Utilisation d'un algorithme d'aide au soin par le patient lui-même :
Consentement du patient

Si aucune exception applicable

Réalisation d'une formalité auprès de la CNIL

2. La phase de déploiement

Cas d'usage : génération de cohortes de patients virtuels dans le cadre d'essais cliniques en cours

- Prérequis: validation de la preuve de concept
- Réflexions en cours :
  - Intégration dans le traitement de données « principal » (l'essai clinique)
  - Modalités d'évaluation éthique
  - Modalités d'information des patients fournissant les données d'entrée (bras contrôle) et participant au bras exposition(quelle explicabilité?)
  - *etc.*

CNIL.

Comment mettre en œuvre le RGPD dans la pratique ?

LES OUTILS POUR VOUS AIDER

Manon de Fallois, adjointe à la cheffe du service de la santé (CNIL)

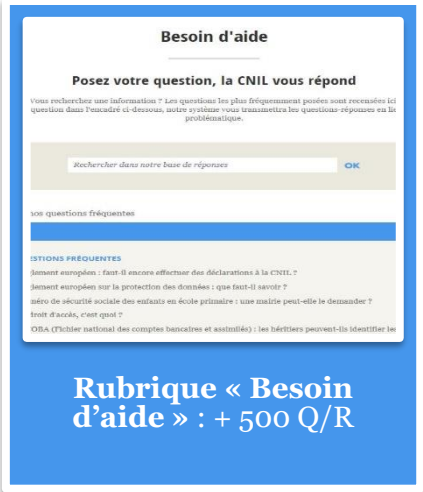
Pour se former

Formation en ligne « L'atelier RGPD »

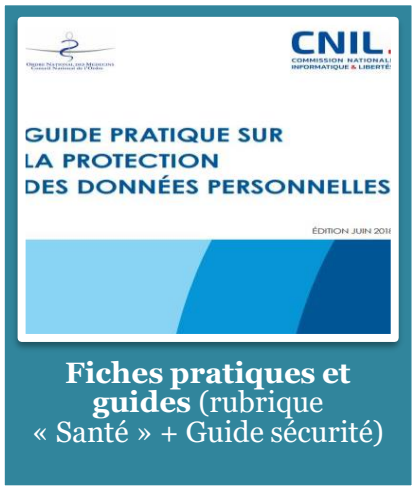
Guides et fiches thématiques

Ateliers

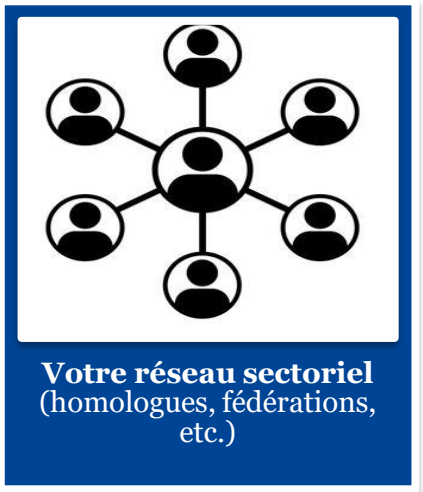
Pour se renseigner



Rubrique « Besoin d'aide » : + 500 Q/R



Fiches pratiques et guides (rubrique « Santé » + Guide sécurité)



Votre réseau sectoriel (homologues, fédérations, etc.)

Les références utiles (1)

- **Les référentiels relatifs à des traitements soumis à autorisation (déclaration de conformité) :**

➢Méthodologies de référence (recherche médicale - MR 001 à MR 006) ; MR-007 et MR-008
➢Référentiel vigilances sanitaires
➢Référentiel entrepôts de données de santé
➢Référentiel accès précoce
➢Référentiel accès compassionnel
➢Référentiel ESND

Les références utiles (2)

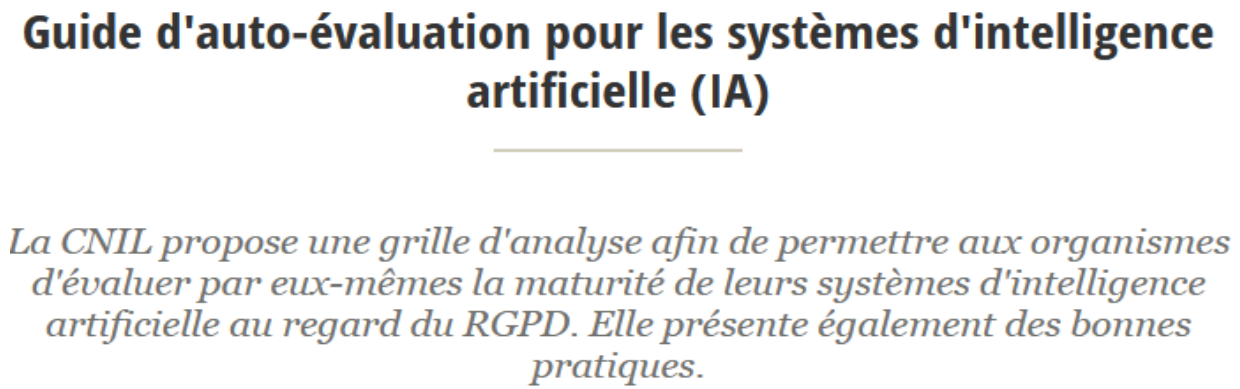
- **Les fiches pratiques:**

➢ **Demandes d'autorisation en santé : la CNIL publie les critères à respecter**
➢ Référentiel des durées de conservation dans le domaine de la santé hors recherche
➢ Référentiel des durées de conservation dans le domaine de la recherche en santé
➢ Référentiel pour la gestion des traitements courants des cabinets médicaux et paramédicaux
➢ Référentiel sur la gestion des officines de pharmacie

➢ Guide sur les modalités de circulation du NIR pour la recherche en santé aux fins d'appariement de données avec le SNDS
➢ Guide pratique sur les durées de conservation
➢ Guide du sous-traitant
➢ Guide pratique sur les mesures de sécurité élémentaires à mettre en œuvre

Pour en savoir plus

- Voir la rubrique « IA » disponible sur le site web de la CNIL:
  - « Petit glossaire de l'intelligence artificielle (IA) »
  - « Professionnels, comment se mettre en conformité ? »
  - « Bac à sable » intelligence artificielle et services publics : la CNIL accompagne 8 projets innovants »



Pour poser une question

Par téléphone au 01 53 73 22 22

Permanence juridique

- Du lundi au vendredi (sauf le mercredi)
- De 10h à 12h et de 14h à 16h

Permanence santé

- Le lundi de 9h30 à 12h.

Permanence DPO

- Du lundi au vendredi (sauf le mercredi)
- De 10h à 12h

Merci de votre attention !