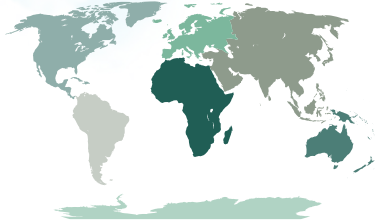




# Fiche récapitulative : Données de santé et qualité des données

## LE PROBLÈME :



1. Enregistrement des informations importantes (recensement naissance/décès, info santé + population)
2. Données fragmentaires (*petits bouts de données*)
3. Pas de statistiques fiables sur des indicateurs important (mortalité, décès, naissances)
4. La moitié des pays ne disposent pas de données sur les causes de décès en 2014 (OMS),
5. + de 100 pays ne disposent pas de systèmes fiables pour enregistrer les naissances et les décès.
6. De nombreux pays = **pas de données de qualité** sur leurs personnels de santé/système de financement de la santé

**Résumé = la présence de données de qualité n'est pas valable pour l'ensemble des pays du monde ON A UN GROS SOUCIS DE QUALITÉ !**



## L'OBJECTIF :

**AVOIR DES DONNÉES DE QUALITÉS DANS UN SYSTÈME DE SANTÉ EST  
NÉCESSAIRE POUR AMÉLIORER LA SANTÉ**

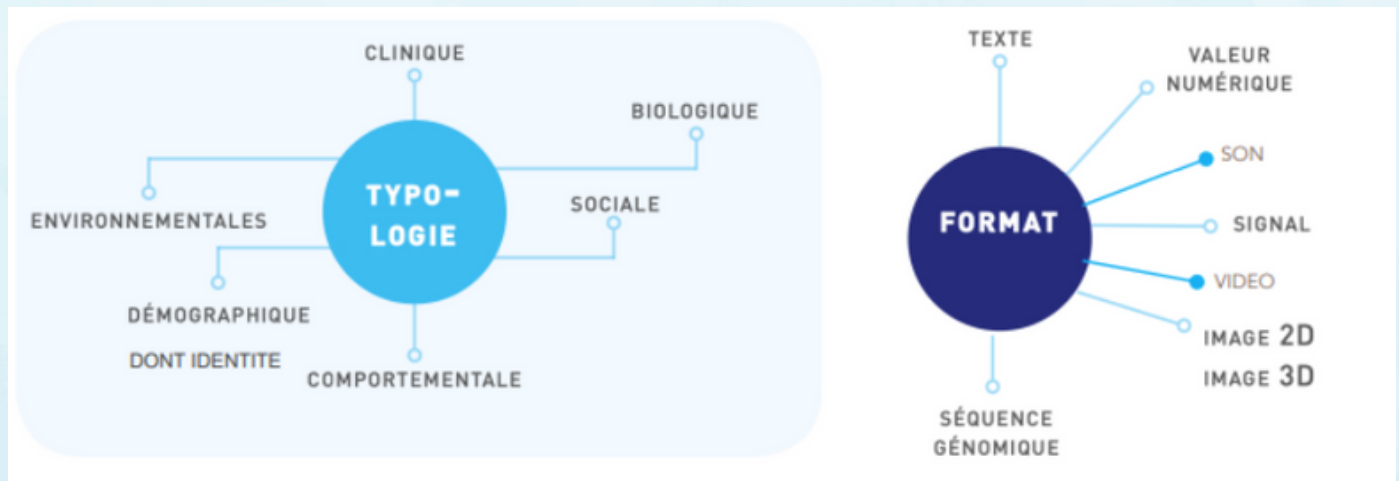




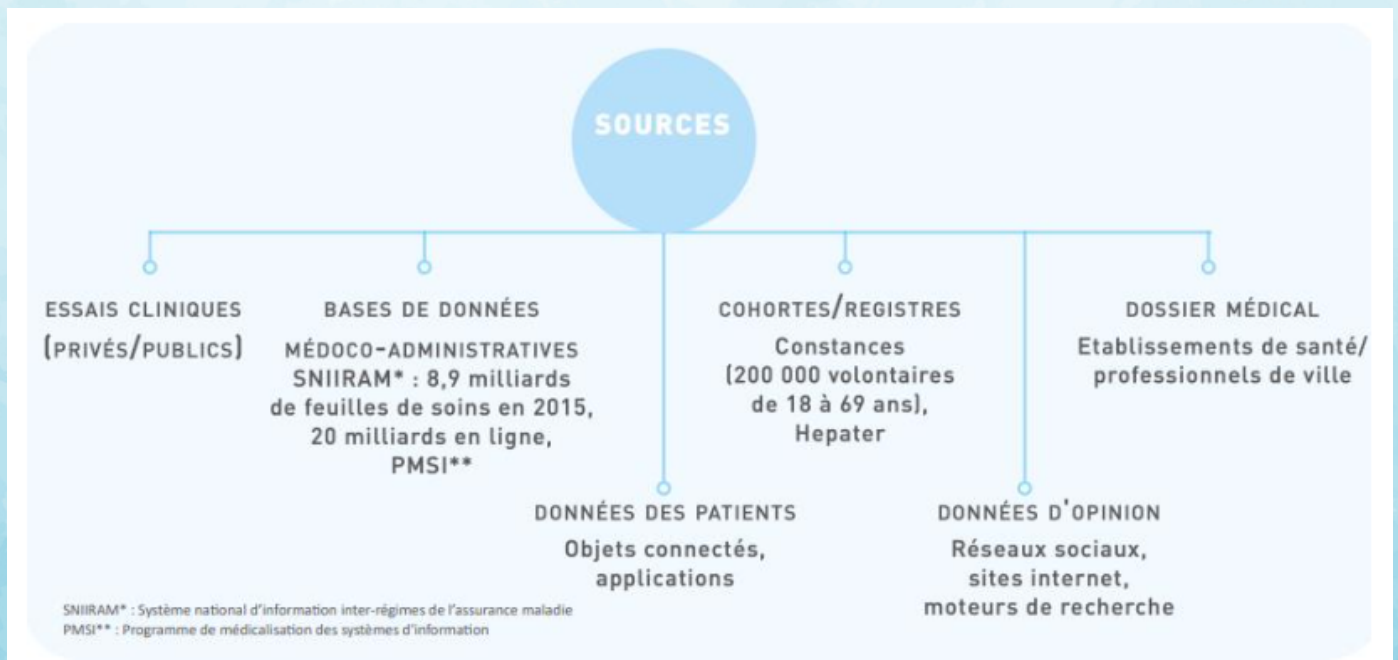
**Données de santé** : Elles régies par la loi informatique et libertés (loi IFL du 6 janvier 1978) définissant les "données à caractère personnel" et la jurisprudence.

**OBJECTIF du règlement Européen sur la Protection des Données, applicable depuis mai 2018** : Renforcer les droits des personnes + responsabiliser les acteurs autour des données, et en particulier les données de santé, qui ont leur définition propre, soit des **"données à caractère personnel relatives à la santé physique ou mentale d'une personne physique, y compris la prestation de service de soins de santé, qui révèle des informations sur l'état de santé de cette personne"**.

En gros on a plein de types de données sur différents aspects de la personne mais aussi plein de format des données !



**On a PLEIN DE SOURCES DIFFÉRENTES DES DONNÉES**





## La source de données

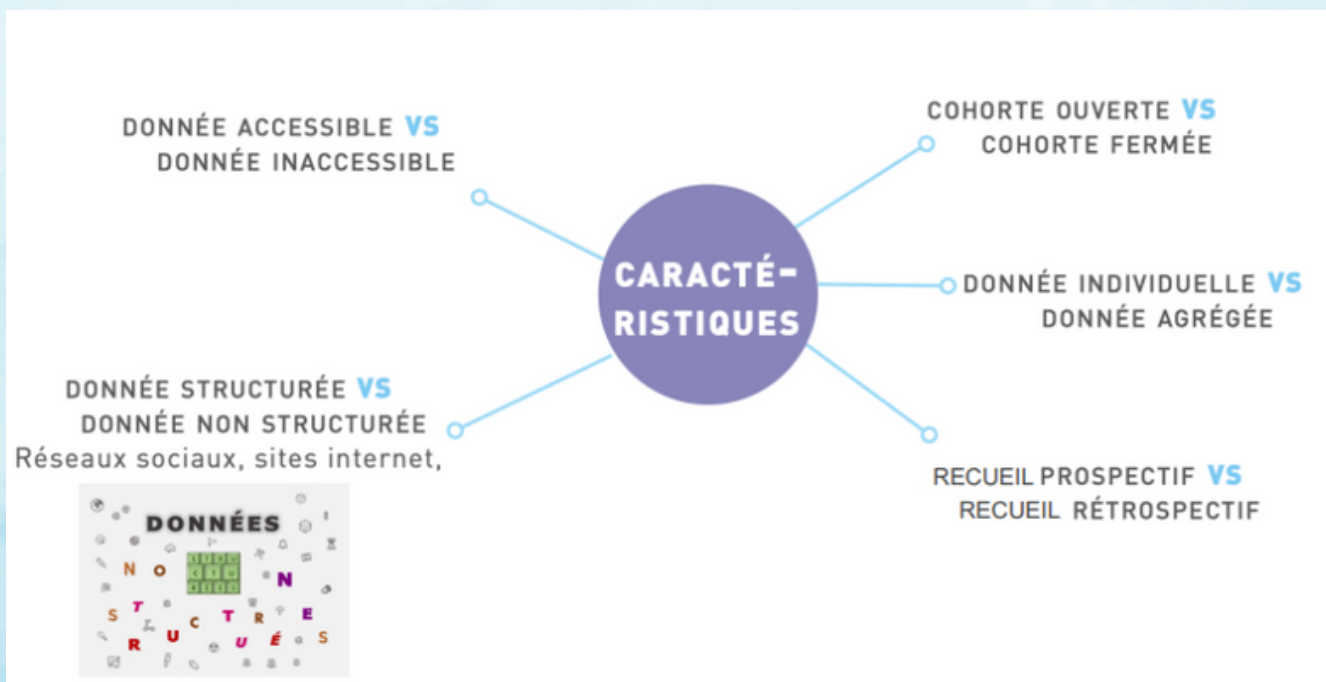
Elles sont le cadre opérationnel de la production. Elles décrivent plus l'**environnement de stockage** et de **mise à disposition** pour différents usages. On trouve dans les sources de données : des données individuelles, collectives ou agrégées.

## Le producteur

Le producteur produit la donnée : c'est un professionnel de santé, c'est aussi possiblement une machine.

D'autres distinctions doivent être faites, qui portent sur l'origine des données et sur l'objectif de leur collecte. Il y a essentiellement **3 types de sources** de ce point de vue :

- Données recueillies en vue de la **gestion et du financement** (en général dans des bases exhaustives)
- Données recueillies pour le **soin** (les dossiers médicaux des établissements et des professionnels, jusqu'au Dossier médical personnel)
- Données recueillies par enquête pour la **santé publique, la veille sanitaire, la recherche épidémiologique**.



❖ **Données structurées** = informations (mots, signes, chiffres...) contrôlées par des référentiels et présentées dans des cases (les champs d'une base de données) qui permettent leur interprétation et leur traitement par des machines.

❖ **Données non-structurées** = le reste, tout ce qui n'est pas organisé en base de données, c'est-à-dire la bureautique, la messagerie, les images, les vidéos, etc.



« Données de vie réelle », ou « données de vraie vie » : Données qui sont générées à l'occasion des soins réalisés en routine pour un patient, et qui reflètent donc a priori la pratique courante.

Les catégories de données relatives à la santé sont les suivantes :

- **Données personnelles sur les citoyens / patients.** Même si des moyens techniques spécifiques sont utilisés (hébergeurs agréés pour les données partagées, numéro d'identification spécifique, carte de professionnel de santé, etc.), les données relèvent du droit commun pour la protection des données individuelles et sont donc sous le contrôle de la CNIL.

→ *Les techniques sont de la responsabilité principalement de l'Agence du Numérique en santé (ancienne Agence des systèmes d'information partagée de santé - ASIP Santé), de la Caisse nationale d'assurance maladie des travailleurs salariés et du GIE SESAM-Vitale.*

- **Données agrégées, statistiques épidémiologiques etc.** qui résultent toujours de traitements de données individuelles collectées pour la gestion ou pour des enquêtes et études spéciales.

- **Données sur l'offre – caractéristiques et activité des hôpitaux, tarifs de professionnels etc.** Celles-ci approchent une autre problématique, fréquente pour les données publiques : la protection de l'information sur l'entreprise.

## Qualité des données

**La qualité des données repose sur leur fiabilité.**

Lorsqu'on dispose de données actualisées de grande qualité, on peut les utiliser en toute confiance dans le cadre :

- Du processus de soins (traçabilité, suivi et continuité des soins, démarche décisionnelle)
- De travaux de recherches utiles
- De planification stratégique pertinente
- De gestion de la délivrance des soins de santé.



**Les critères : ++++**

- la pertinence et la complétude
- la fiabilité
- la validité, la cohérence
- l'exactitude, la précision
- l'actualité, la régularité
- la compréhension, l'intelligibilité
- l'accessibilité



<i>Pertinence</i>	Les données répondent à la question posée
<i>Complétude</i>	Valeurs manquantes qui rend les données <b>incomplètes</b> . <b>Solution</b> = utilisation d'une bonne source de données
<i>Fiabilité</i>	Les données correspondent à la situation réelle
<i>Validité</i>	<b>Degré de conformité</b> des données aux règles ou contraintes définies (plage, obligatoires, unicité, motifs d'expression régulière, clé étrangère, validation entre champs)
<i>Cohérence</i>	<b>Validation interne de la base de données</b> , mais aussi et surtout à la comparabilité des données et des résultats à des connaissances antérieures
<i>Exactitude et précision</i>	On doit vérifier la précision des données et pas seulement leur exactitude
<i>Actualité</i>	Elle est aussi appelée aussi <b>récence</b> . C'est le temps qui s'écoule entre la collecte des données et la parution des résultats.
<i>Régularité</i>	Les données doivent être collectées avec rapidité (sondages uniques) ou une fréquence (collecte régulière de données, telle que les registres) convenant à l'usage prévu.
<i>Compréhension et intelligibilité</i>	Les données doivent être collectées en utilisant la terminologie technique standard. - Il est également important de les retranscrire dans un langage que les utilisateurs ciblés comprennent. - L'intelligibilité porte sur la documentation de la méthodologie employée, sa clarté, sa compréhension par des utilisateurs non-spécialistes.
<i>Accessibilité</i>	Mode de restitution, de mise à disposition, de présentation. La visualisation en fait partie.