

Panorama des données de santé en France

Dominique Polton

23/09/2023

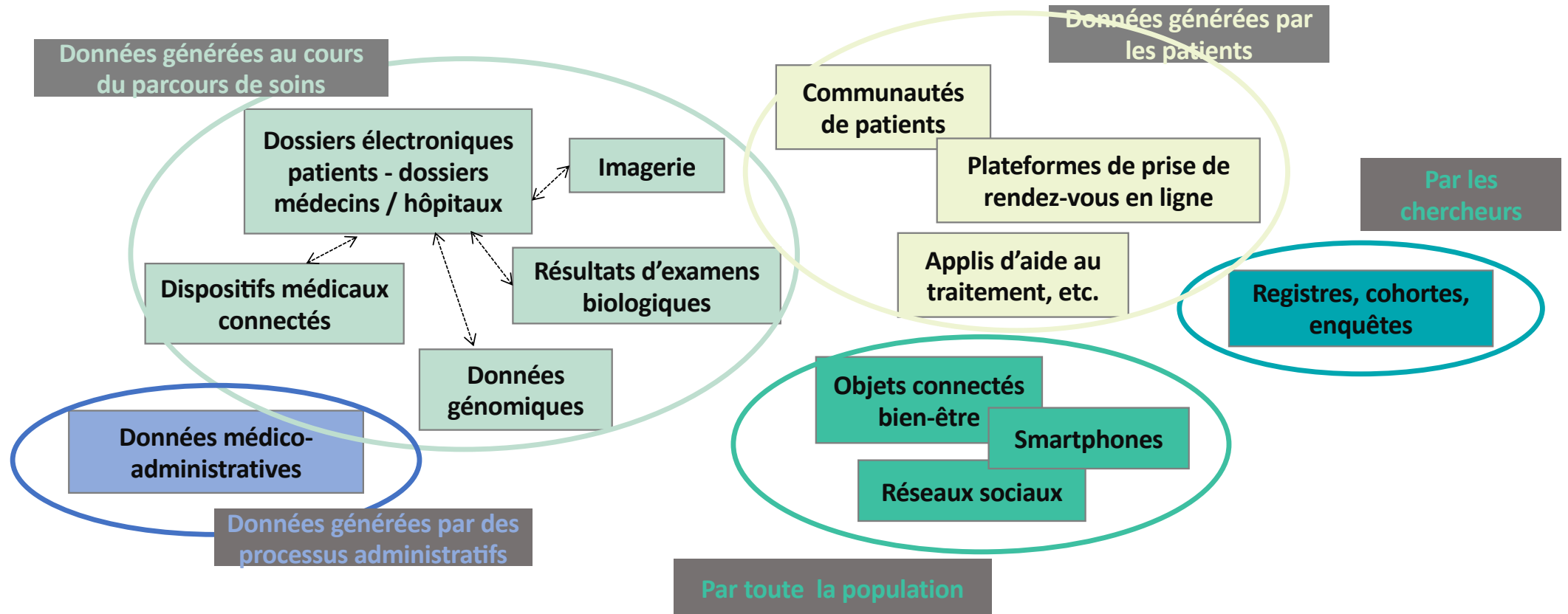
1. Les données de santé c'est quoi ?
2. La construction d'un patrimoine des données de santé en France : l'évolution des 20 dernières années
3. Les enjeux aujourd'hui pour développer l'usage des données de santé au bénéfice de tous

Les données de santé, c'est quoi ? – une définition récente

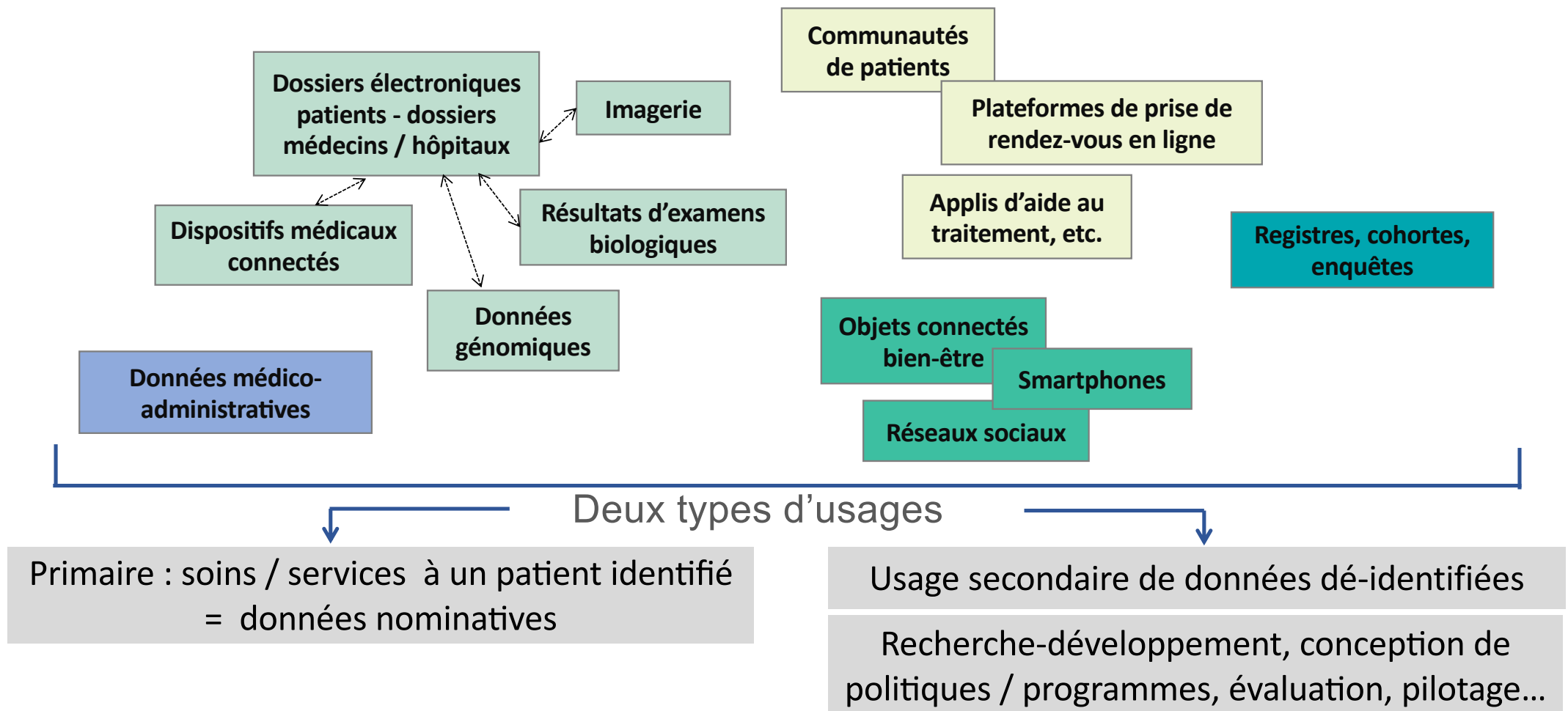
- Encadrement juridique ancien
 - ✓ loi informatique et liberté de 1978 → **protection de la vie privée**
 - ✓ directive européenne de 1995 sur la protection des données à caractère personnel → « données sensibles » (interdiction de collecter /traiter des données [*faisant apparaître*] les origines raciales ou ethniques, les opinions politiques, philosophiques, religieuses, l'appartenance syndicale, ou relatives **à la santé** ou à la vie sexuelle)

- Mais finalement pas définies explicitement jusqu'au Règlement général sur la protection des données (RGPD – 2018), qui en donne une **définition très large**

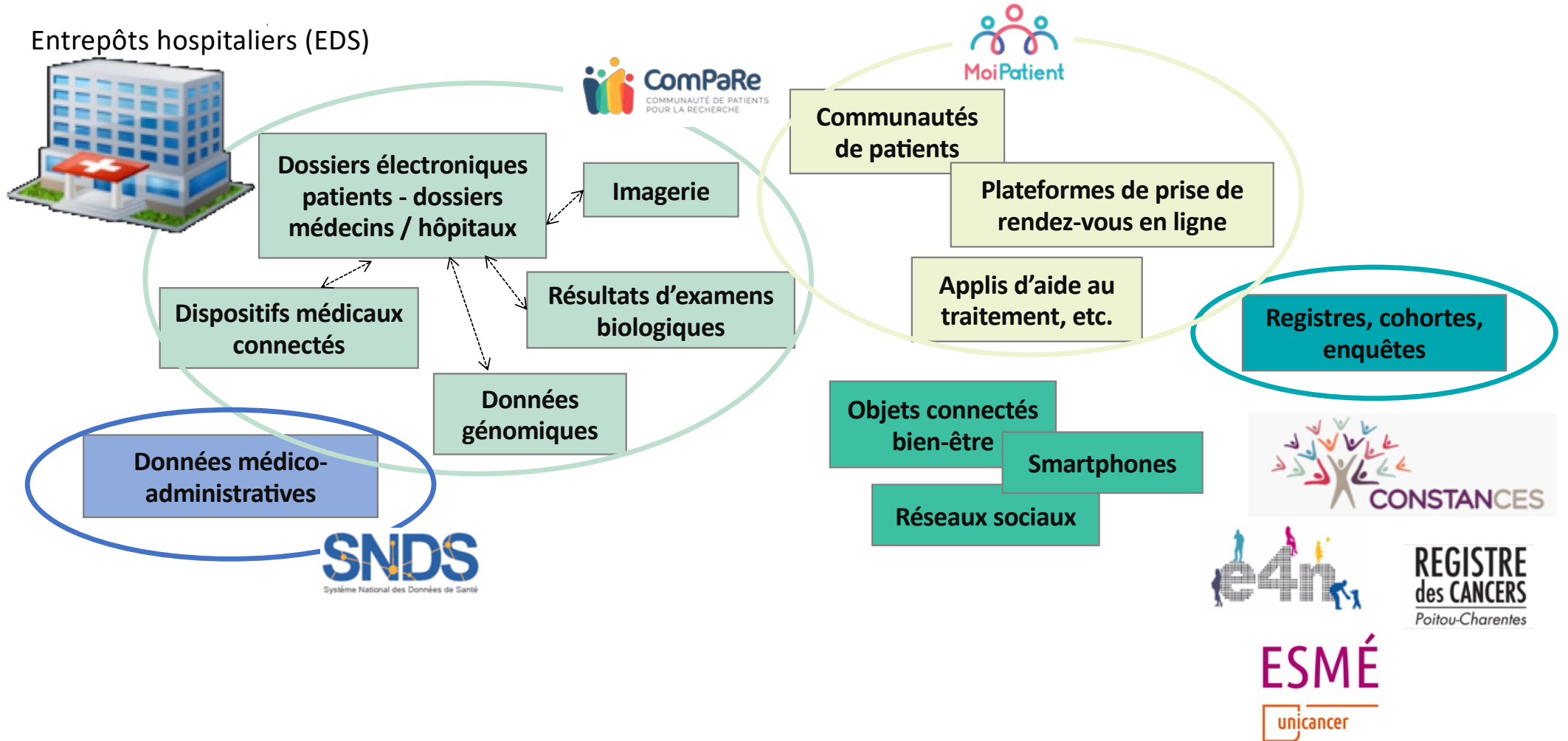
Les données de santé, c'est quoi ? – des sources multiples



Les données de santé, c'est quoi ? Deux usages à distinguer



Bases de données pour des usages secondaires, quelques exemples



1. Les données de santé c'est quoi ?
2. La construction d'un patrimoine des données de santé en France : l'évolution des 20 dernières années
3. Les enjeux aujourd'hui pour développer l'usage des données de santé au bénéfice de tous

Un chemin spécifique par rapport à d'autres pays

- Certains pays sont partis de **bases de données cliniques** constituées depuis longtemps (et ont eu une avance considérable de ce point de vue)

Exemple UK : GPRD (General Practice Research Database (devenu CPRD), créé en 1987 - échantillon de dossiers médicaux de généralistes, 6 millions de patients, depuis 1987 – large accès, utilisation ++ en pharmaco-épidémiologie (1000 articles publiés dans des revues scientifiques)

Exemple pays nordiques : nombreux registres de population nationaux (dont des registres cliniques) & numéro d'identification unique permettant les croisements entre bases de données → production de multiples études en épidémiologie, évaluation des traitements et qualité des soins

- La France a pris un autre point de départ, mais le patrimoine de données que nous avons constitué, avec une **base de données médico-administrative** unique en Europe, a un potentiel énorme si nous relevons collectivement les enjeux

L'évolution sur les 20 dernières années



- 2010 - chaînage SNIIRAM-PMSI en routine
- 2005 - Entrepôt opérationnel – progressivement enrichi
- LFSS 1998 - création du SNIIRAM

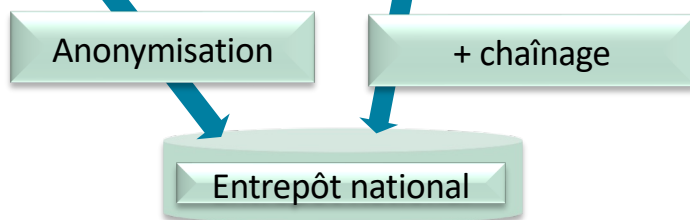
Systeme national inter-régimes de l'assurance maladie (SNIIRAM) : des feuilles de soins numérisées, anonymisées et centralisées

Feuilles de soins



2009-2010 : industrialisation du chaînage avec le PMSI

Factures des hôpitaux



Des données limitées (pas de données cliniques & paracliniques, de facteurs de risque, peu de données sociales...)

MAIS

Une vision complète du parcours de soins du patient,

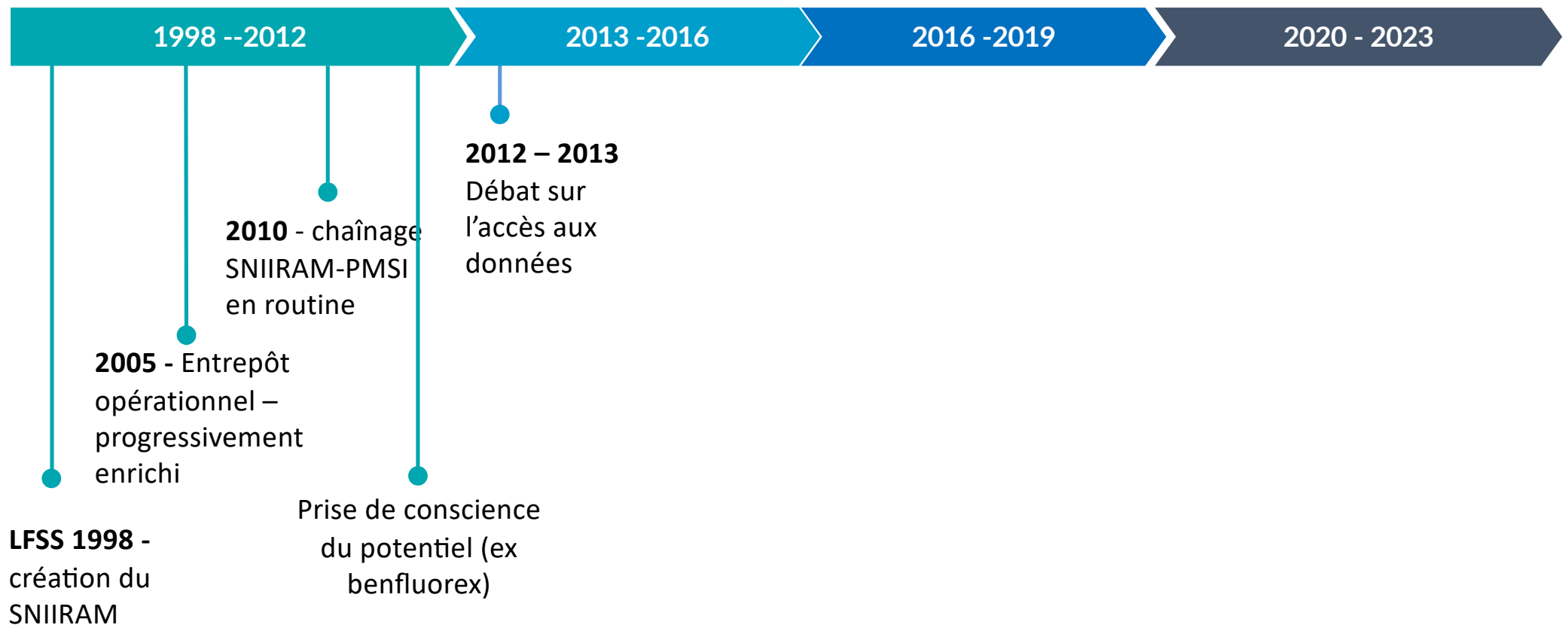
68 millions de personnes, population exhaustive sans biais de sélection,

pas de perdus de vue, des années d'historique, une bonne homogénéité du codage



Nombreuses exploitations possibles

L'évolution sur les 20 dernières années



Plus grave que le débat sur l'affaire des données de santé publique

Les médecins français n'ont pas la culture de santé publique
de Kervasdoué, professeur au CNAM et Didier Sicard, président
du Comité consultatif national d'éthique.

Parce que certains actes rémunérateurs pour le praticien ou l'établissement
hospitalier continuent d'être remboursés, alors que leur justification
clinique a disparu ; parce que l'assurance-maladie non seulement n'exploite
pas ou peu les données qu'elle possède mais aussi et surtout rend difficile
aux chercheurs d'y avoir accès et l'interdit à tous les autres acteurs, même
quand l'autorise la Commission nationale de l'informatique et des libertés !

Une seule base de données est en pratique accessible, celle qui livre des
informations sur tous les séjours hospitaliers de tous les Français une
année donnée, mais elle est tenue par l'Etat, pas par l'assurance-maladie.

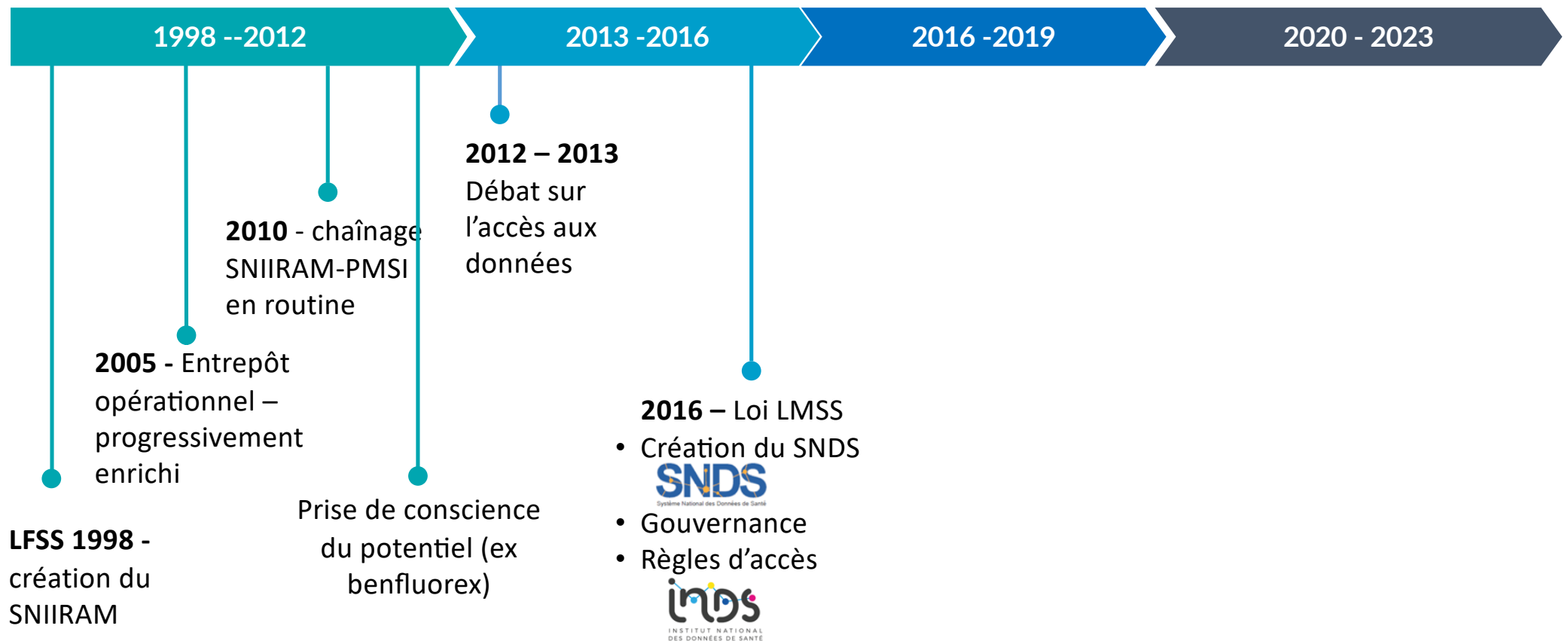
... sous des prétextes fallacieux, empêche de documenter les
inégalités de santé. Selon l'endroit
clinique, compte tenu du

Libérez les données de santé !

07 FEVRIER 2013

C'est le slogan d'Initiative transparence santé, un collectif composé de chercheurs, de
représentants d'usagers et de patients (Que Choisir, CISS), de complémentaires santé,
d'entrepreneurs, de consultants et de journalistes. Le collectif a lancé en ligne une pétition
« Libérez les données de santé ! ».

L'évolution sur les 20 dernières années



Systeme national inter-régimes de l'assurance maladie (SNIIRAM) : des feuilles de soins numérisées, anonymisées et centralisées

2009-2010 : industrialisation du chaînage avec le PMSI

2016 : ajout de la base de données sur les causes médicales de décès

Feuilles de soins



Factures des hôpitaux



Certificats de décès



Anonymisation

+ chaînage

Entrepôt national

A venir

Données sur le handicap

Echantillon ass. complémentaire



Gouvernance unifiée des trois bases, mêmes règles d'accès

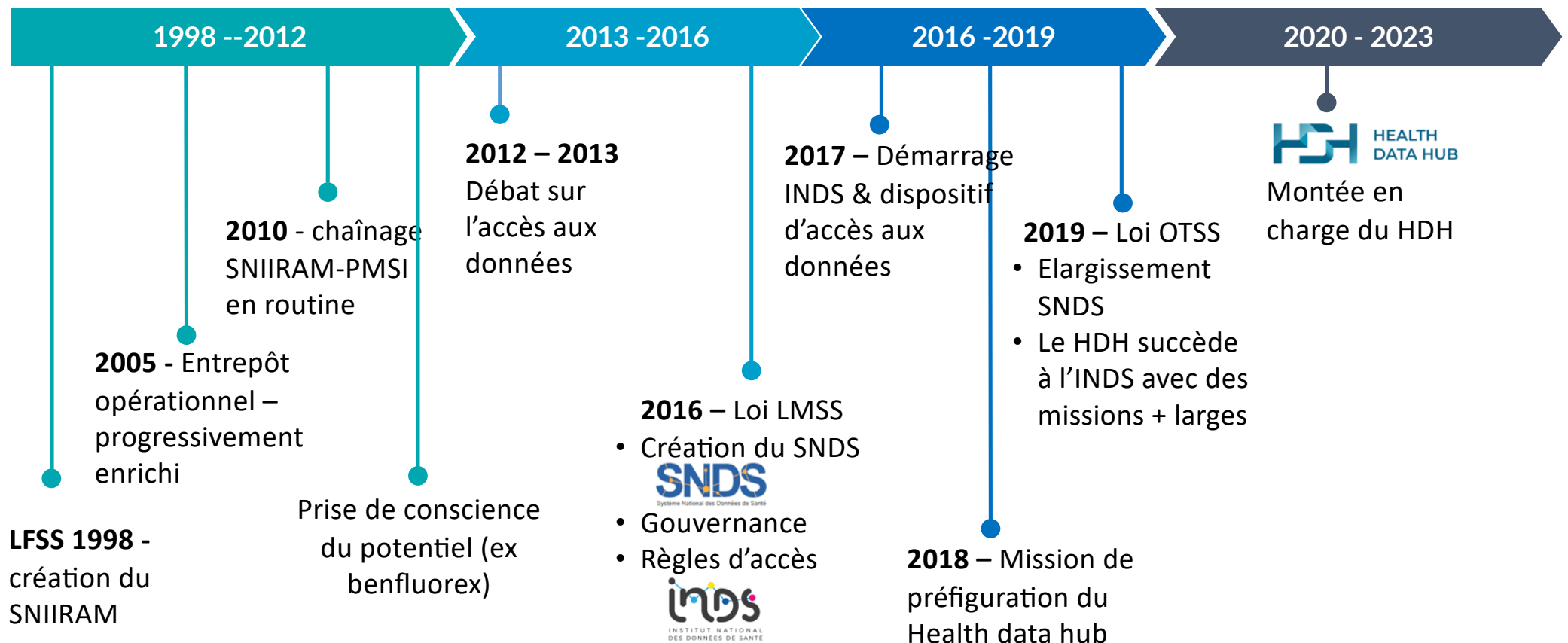
Ouverture aux acteurs publics et privés

- ✓ Accès permanent si mission de service public
- ✓ L'accès sur projet pour tous
- + open data pour les jeux de données ne présentant aucun risque de réidentification

Conditions

- ✓ Intérêt public – finalités interdites
- ✓ Protection de la vie privée – Sécurité
- ✓ Transparence

L'évolution sur les 20 dernières années



→ Les missions du Health data hub

- **Guichet unique** pour accompagner les porteurs de projets dans leurs démarches administratives d'accès aux données
- Elaboration et enrichissement progressif d'un **catalogue de bases de données** partenaires mises à disposition de la communauté scientifique
- Fourniture d'une **plateforme sécurisée à l'état de l'art** pour stocker et traiter les données
- **Accompagnement** de nombreux projets, **services** aux gestionnaires et utilisateurs de données, animation de la communauté, partage d'expériences et de connaissances

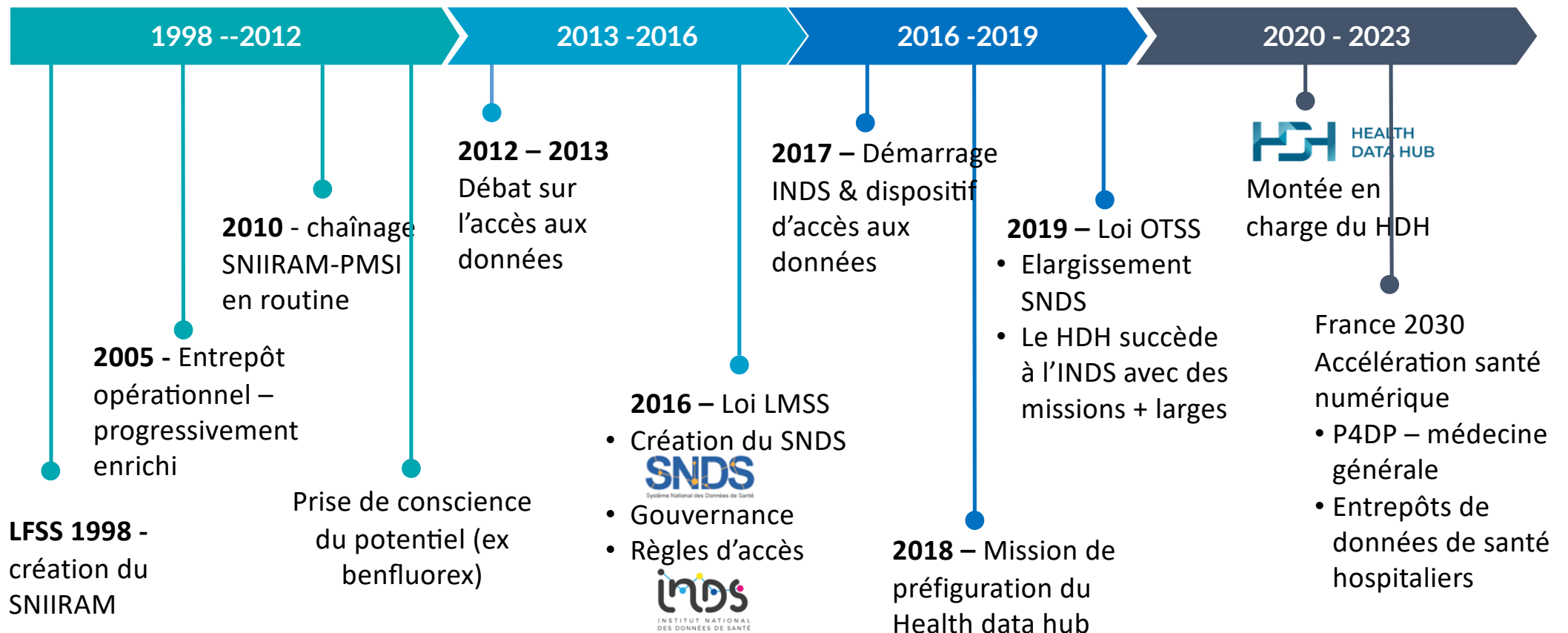
→ Parmi les orientations stratégiques

Co-construire une culture de la donnée de santé avec les citoyens / usagers

Renforcer la place de la France à l'international

En juillet 2022, un consortium piloté par le HDH a été lauréat d'un appel à candidatures de la Commission européenne pour construire une version test du futur espace européen des données de santé.

L'évolution sur les 20 dernières années



1. Les données de santé c'est quoi ?
2. La construction d'un patrimoine des données de santé en France : l'évolution des 20 dernières années
3. Les enjeux aujourd'hui pour développer l'usage des données de santé au bénéfice de tous

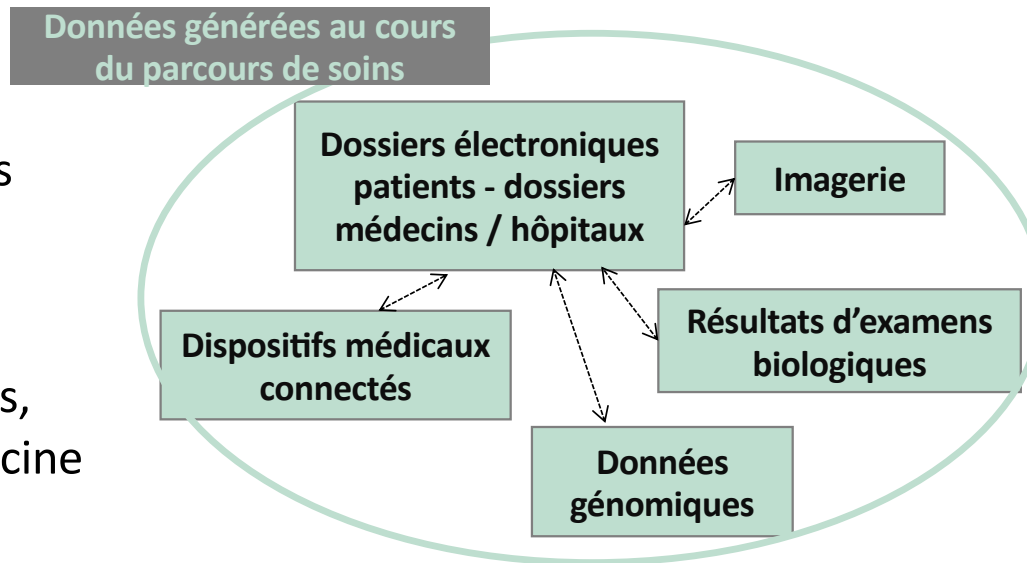
Les enjeux aujourd'hui

1. Il faut des données de qualité, exploitables

La partie faible de la France = données cliniques et paracliniques, récoltées au cours du parcours de soins



Priorités = entrepôts hospitaliers, entrepôts de données de médecine de ville (P4DP) - registres ?



Les enjeux aujourd'hui

2. Il faut pouvoir utiliser ces données (donc y accéder)

- Construire un cadre de confiance pour les citoyens et les patients (transparence, partage des résultats, communication, outils de dialogue)
- Au-delà du cadre législatif, développer une culture du partage (réflexe « propriétaire »)
- Gouvernance des bases de données, procédures d'accès



Aujourd'hui les données sont accessibles à la recherche en théorie, mais les procédures restent complexes, les délais très longs, « un parcours d'obstacles réglementaires, techniques, et administratifs » (Acteurs Publics 2022).



Délai moyen d'accès
aux données du SNDS

Commission des
AFFAIRES SOCIALES



12 JUILLET 2023

B. UNE UTILISATION SECONDAIRE DES DONNÉES ENCORE TROP COMPLEXE

1. Un parcours d'accès qui présente encore des marges de fluidification

Les enjeux aujourd'hui

3. Il faut pouvoir apparier différentes bases et créer des systèmes multi-sources, de façon pérenne ou en fonctions des projets

La richesse est dans le croisement des bases, notamment données administratives avec d'autres sources : exemple SNDS + entrepôts hospitaliers, SNDS + échantillon démographique permanent, SNDS + OSCOUR

Là aussi, la situation reste complexe aujourd'hui pour les chercheurs qui veulent pouvoir croiser les données.

Les enjeux aujourd'hui

4. Il faut accroître la capacité à les exploiter intelligemment

- Capacités technologiques (plateforme à l'état de l'art)
- Problématique des ressources humaines et des compétences
- Enjeu de mutualisation et de partage au sein de la communauté (variables construites, algorithmes...
 - *Exemple du réseau REDSIAM : réseau d'équipes de recherche et d'agences qui construisent collectivement des algorithmes pour repérer des pathologies à partir des données du SNDS*