



Ministère de l'Economie et des Finances

Ministère des Affaires sociales et de la Santé

Ministère du Travail, de l'Emploi,
de la Formation professionnelle et du Dialogue social

Données de santé

André Loth

Typologie des données de santé (Etalab)

Données socio-démographiques	Santé publique et épidémiologie	<ul style="list-style-type: none"> • Etat de santé • Epidémiologie • Habitudes de vie • Environnement • Infrastructures • Services proposés • Personnel • Tarifs et honoraires
	Offre de soins	<ul style="list-style-type: none"> • Infrastructure de recherche • Informations médicales • Activité des établissements de santé • Consultations • Biologie • Transport • Médicaments et dispositifs médicaux • Indemnités et accidents de travail • Activité des établissements médico-sociaux
	Consommation de soins	<ul style="list-style-type: none"> • Autres • Performance financière • Performance opérationnelle • Qualité • Dépenses de soin et de biens médicaux • Dépenses de soin de longue durée • Dépenses de prévention • Autres dépenses soins • Dépenses administratives • Modalités/sources de financement
	Performance et opérations	
	Autres dépenses	

- Sources : voir rapport 2009 du HCSP sur les données de santé
- Recensement des données disponibles en cours par Etalab
- Mission IGAS sur les informations à la disposition du public (offre de soins, indicateurs de qualité, conduites à tenir...)

Données de santé : vue d'ensemble

- **Données médicales à visée individuelle** (pour soigner) : dossiers médicaux (compte-rendus, prescriptions, images, résultats d'examen, prélèvements, données génétiques...) : très nombreuses, très sensibles...
 - **Données médico-administratives à visée métier** (pour administrer, gérer, contrôler, autoriser payer...) : très nombreuses, très sensibles...
- **Données à visée de connaissance** : souvent les mêmes que les précédentes mais
 - Collectées par des chercheurs, organismes d'études ou services statistiques (enquêtes, panels, cohortes, avec réutilisation de dossiers médicaux, de registres, de biobanques...)
 - Ou recyclées par les administrations sous une forme « déidentifiées »...
 - Puis réutilisées éventuellement par d'autres (chercheurs, institutions, entreprises, journalistes)...
 - sous certaines conditions
 - et sous diverses formes (données granulaires, agrégées, indicateurs et résultats d'analyse etc.)

Données médico-administratives et SNIIR-AM

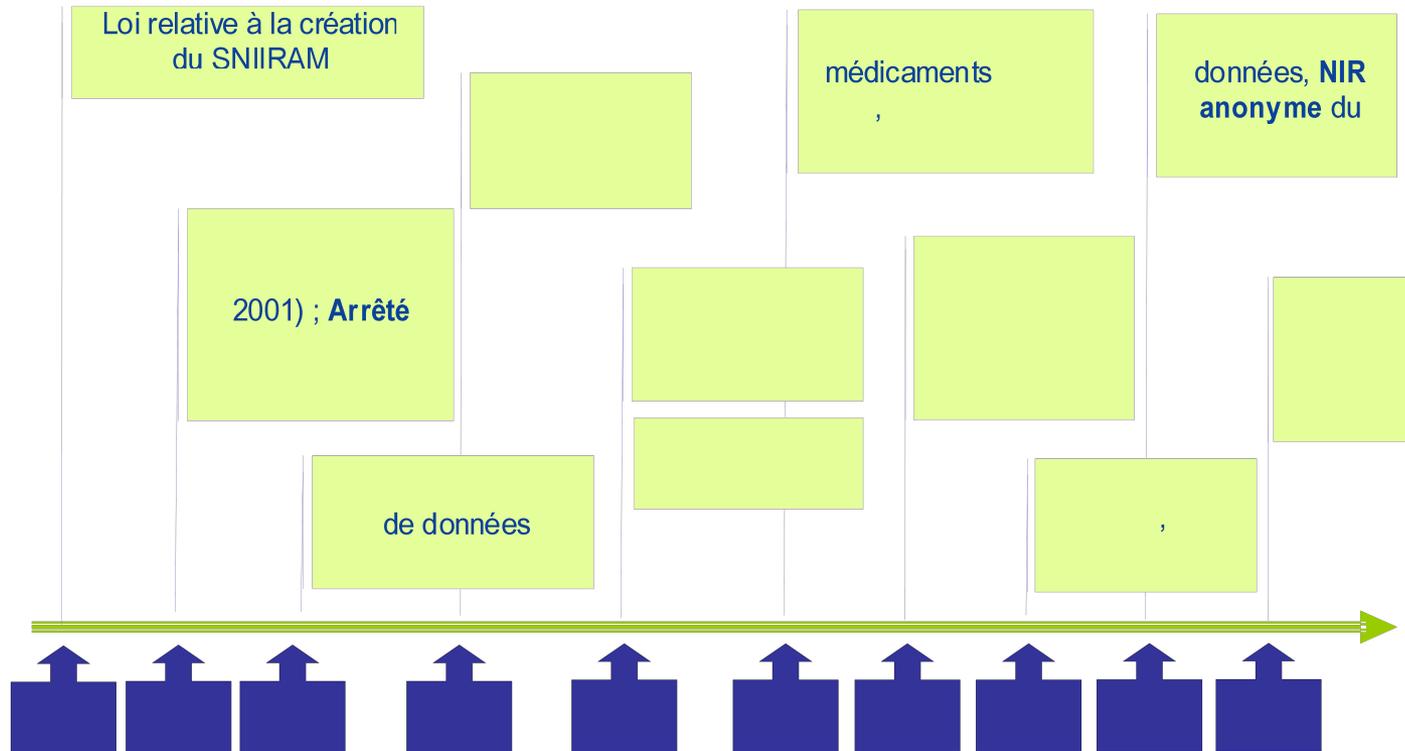
- Données médico-administratives à vocation exhaustives:

- Décomptes assurance maladie (SNIIRAM)
- Résumés de sortie hospitaliers (PMSI)
- Certificats de décès (CépiDc)
- Certificats de santé des enfants...
- Autorisations d'activité (référentiels : FINESS, RPPS...)
- Etc.

SNIIRAM étendu:

- Toute la population
- Age, sexe, commune de résidence, date de décès
- Soins de ville (y c. médicaments)
- Soins hospitaliers
- Arrêts de travail...
- ALD, CMU...
- **Chaînage (suivi longitudinal par individu)**

Une base récente



Une richesse sous-utilisée

- Enjeux démocratiques, sanitaires et économiques
- Obstacles juridiques (complexité, tabou du NIR pour appariements) et obstacles techniques
- Critiques des chercheurs
- L'Institut des données de santé
- Polémique début 2013
- Rapport Bras sur la gouvernance et l'utilisation des données de santé (octobre 2013)

Préconisations du rapport Bras : deux types d'accès

- Données anonymes : Open data
 - Cas particulier des données nominatives sur l'activité des professionnels
- Données indirectement nominatives : accès restreint et traçabilité (modèle CASD : pas de copies) selon projet, pour les chercheurs et les études d'intérêt public (demandeurs publics ou privés)
 - Cas particulier des organismes dont la mission requiert un accès permanent

Quatre chantiers en cours à la suite du rapport Bras dans la perspective de la loi santé

- « Commission open data », pluripartite, co-animée par SG-DSSIS (Ph. Burnel) et DREES (F. von Lennepe) avec l'appui d'Etalab
- Etude technique « Risque de réidentification »
- Travaux juridiques (règles d'accès, NIR...)
- Etude de cadrage (les services, les moyens...) avec l'appui du SGMAP

Zoom 1 : données personnelles, données « déidentifiées », données anonymes

- Feuilles de soins et dossiers médicaux sont nominatifs
- Les données du SNIIRAM-PMSI sont dé-identifiées (+ NIR chiffré de manière irréversible) mais pas forcément anonymes :
 - Si les données sont précises et exhaustives, on y reconnaît des personnes connues (proches, célébrités, employés, clients...)
 - Si par ex. on connaît hôpital, mois de sortie, durée du séjour, date (m/a) de naissance, sexe, code postal, alors 89 % des hospitalisés sont identifiables dans la base PMSI (Blum D., Trouessin G., 2011)
 - *Plus il y a de personnes, dans le secret moins il y a de secret*
- Pas d'open data pour les données *personnelles* de santé !
Prétexte ? Non : droit à la vie privée...
 - L'objectif c'est l'ouverture la plus large possible (maximum de données, maximum d'utilisations) compatible avec la protection de la vie privée

Zoom 2 : gouvernance (chantier en cours !)

- Où s'élaborent et se discutent doctrine, critiques, demandes, évaluation... :

- Quelles données ? Comment les améliorer ? Qui accède à quoi ? Sous quelles conditions ?
- Bilan, améliorations...

Un « comité d'orientation des parties prenantes »

- Qui gère concrètement les données ? CNAMTS, ATIH...
- Qui donne concrètement l'accès (pour un projet...) ?
 - CNIL
 - et en amont : comités de protection des personnes (les CPP et leur comité national d'harmonisation), le CCTIRS (comité d'experts), le comité d'orientation...
- Etat plus présent et plus attentif aux demandes