

Commission open data en santé

Annexes du RAPPORT

Juillet 2014

Annexes

Annexe 1 : Composition de la commission open data

Annexe 2 : Lettre de mission

Annexes 3a et 3b : politique du gouvernement français en matière d'open data (feuille de route du gouvernement, décisions prises en CIMAP)

Annexe 4 et 4 bis : compléments sur l'état des lieux juridique

Annexe 5 et 5 bis : compléments sur les accès au SNIIRAM et au PMSI

Annexe 6 : Cartographie des données de santé

Annexe 7 : Panorama international et principaux enseignements

Annexe 8 : Compte-rendu des ateliers thématiques

Annexe 8 bis : Synthèse de la consultation publique

Annexe 9 : rapport du groupe technique « risques de réidentification »

Annexe 10 : Contributions des membres

Annexe 1
Composition de la commission

Philippe	BURNEL	DSSIS	Délégué à la stratégie des systèmes d'information de santé
Christian	BABUSIAUX	IDS	Président
Christian	BAUDELLOT	ENS	Sociologue
Max	BENSADON	ATIH	Directeur adjoint
Kat	BORLOGAN	Five by Five	Co-fondatrice
Yann	BOURGUEIL	IRDES	Directeur
Chantal	CASES	INED	Directrice
David	CASTILLO	FHP	Responsable des études économiques et des SI
Brigitte	CONGARD-CHASSOL	SNITEM	Directeur Affaires Médicales, Economiques et Export, Responsable Qualité
Patrick	CORNE	UNPS	Masseur-Kinésithérapeute
Mathieu	COUSINEAU	UNOCAM	FNMF - responsable pôle assurance santé
Laurent	CYTERMANN	Conseil d'Etat	Rapporteur général adjoint
Bénédicte	DELAUNAY	Université de Tours	Professeur de droit public, membre de la CADA
Mathieu	ESCOT	UFC QUE CHOISIR	Chargé de mission santé
Martial	FRAYSSE	CNOP	Président du conseil régional des pharmaciens d'Ile-de-France
Yves	GAUBERT	FHF	Responsable pôle Finances
Pierre-Yves	GEOFFARD	PSE - ENS	Directeur de l'Ecole d'Economie de Paris
Claude	GISSOT	CNAMTS	Directeur de la stratégie, des études et des statistiques
Jean-François	GOGLIN	FEHAP	Conseiller Systèmes d'Information de Santé
Marcel	GOLDBERG	INSERM	Chercheur
Evelyne	GUILLET	UNOCAM	CTIP - directrice santé
Chrystel	JOUAN-FLAHAULT	LEEM	Directeur médical du Leem
William	JOUBERT	UNPS	Médecin - vice-président UNPS
Pierre	LEPORTIER	UNPS	Pharmacien - vice-président UNPS
Frédérique	LESALNIER	CNIL	Responsable pôle santé
André	LOTH	DREES	Directeur de projet
Jacques	LUCAS	CNOM	Vice-président
Bertrand	LUKACS	APHP	Chirurgien urologue
Cécile	MALGUID	UNOCAM	FFSA - responsable des études santé
Odile	MARQUESTAUT	FNEHAD	-
Jacques	MARZIN	DISIC (SGMAP)	Directeur
Jean-Roch	MEUNIER	MEDICEN PARIS REGION	Délégué Général
Jean-Paul	MOATTI	INSERM ; Aviesan	Directeur ITMO santé publique
Michel	PINAULT	COEPIA	Président
Philippe	RANQUET	DAJ	directeur des affaires juridiques des Ministères sociaux
Grégoire	REY	CépiDc-INSERM	Directeur

Emmanuel	REYRAT	FNCLCC	Directeur du département des systèmes d'information
Christophe	RICHARD	SYNTEC	Medical director Santeos
Christian	SAOUT	CISS	Président d'honneur du CISS
Bruno	SARFATI	SYNTEC	Vice-Président Cegedim
Henri	VERDIER	Etalab (SGMAP)	Directeur
Franck	VON LENNEP	DREES	Directeur de la recherche, des études, de l'évaluation et des statistiques
Mahmoud	ZUREIK	ANSM	Directeur de la stratégie et des affaires internationales

Annexe 2

Lettre de mission



[Lettre de mission.pdf](#)

Annexe 3 : Politique du gouvernement français en matière d'open data :

Feuille de route du gouvernement français sur l'ouverture et le partage des données publiques (3a), extraits des relevés des décisions prises sur l'open data en Comité Interministériel pour la Modernisation de l'Action Publique (3b), et Vademecum

Annexe 3a – Feuille de route du gouvernement français sur l'ouverture et le partage des données publiques

L'ouverture des données publiques, liberté publique et levier d'innovation

L'ouverture des données publiques est fondée sur la loi n° **78-753** du 17 juillet 1978 relative au droit d'accès aux documents administratifs, ainsi définis : « (...) quels que soient leur date, leur lieu de conservation, leur forme et leur support, les documents produits ou reçus, dans le cadre de leur mission de service public, par l'Etat, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission. Constituent de tels documents notamment les dossiers, rapports, études, comptes rendus, procès-verbaux, statistiques, directives, instructions, circulaires, notes et réponses ministérielles, correspondances, avis, prévisions et décisions. (...) » (article 1^{er}).

Le droit d'accès aux documents administratifs a été reconnu comme une « liberté publique » par le Conseil d'État (CE, 29 avril 2002, U., n° 228830). En 1997, le gouvernement en a élargi le principe en décidant la mise en ligne gratuite des « données publiques essentielles ». En 2003, la directive 2003/98/CE du Parlement européen et du Conseil du 17 novembre 2003 concernant la réutilisation des informations du secteur public, transposée par l'ordonnance du 6 juin 2005 et le décret du 30 décembre 2005, a facilité la réutilisation des documents existants détenus par des organismes du secteur public. Le décret n° **2011-577 du 26 mai 2011 relatif à la réutilisation des informations publiques détenues par l'Etat et ses établissements publics administratifs** a posé le principe de la gratuité du droit à réutilisation.

Le 21 février 2011, le décret n° 2011-194 portant création d'une mission « Etalab » chargée de la création d'un portail unique interministériel des données publiques créait la structure en charge d'animer et d'héberger la diffusion de ces données publiques.

Avec la révolution numérique - qui facilite la production, la diffusion et le traitement de données de masse - l'ouverture et le partage des données publiques sont devenus un puissant levier pour :

- renforcer la confiance des citoyens grâce à une plus grande transparence de l'action publique (par exemple, avec la mise en ligne d'informations sur les comptes publics, la pollution ou les délits constatés) ;
- permettre de nouvelles formes de co-production avec la société et soutenir l'innovation sociale (comme le projet Handimap qui a permis, grâce aux données des villes de Rennes et de Montpellier, de développer une application de calcul d'itinéraire pour handicapés moteurs) ;
- améliorer le fonctionnement administratif (comme en témoigne la très forte utilisation des données publiques par l'administration elle-même) ;

- améliorer l'efficacité de l'action publique en développant de nouveaux modes d'organisation et de nouveaux processus de travail (comme le suivi des accidents de la route par la sécurité routière permettant d'améliorer l'aménagement de la voirie) ;
- soutenir le dynamisme économique, en créant de nouvelles ressources pour l'innovation et la croissance (comme en témoignent les jeunes entreprises innovantes primées par Dataconnexions) ;
- contribuer à l'influence et au rayonnement international de la France (initiative *Open Government Partnership* -OGP-, initiative sur la transparence de l'aide internationale -IATI- ou sur la transparence des industries extractives - EITI).

L'engagement du Gouvernement en faveur de la transparence de l'action publique et de l'ouverture des données

Dès le premier Conseil des ministres, le 17 mai 2012, le Président de la République a fait signer, à tous les membres du gouvernement, une charte de déontologie rappelant notamment leur « devoir de transparence », leur « scrupuleux respect des dispositions garantissant l'accès des citoyens aux documents administratifs » et leur engagement à « mener une action déterminée pour la mise à disposition gratuite et commode sur Internet d'un grand nombre de données publiques ».

En lançant le chantier de la modernisation de l'action publique, le Premier ministre a réaffirmé le 31 octobre 2012 les principes d'ouverture et de partage des données publiques prévus par la circulaire du Premier ministre du 26 mai 2011, notamment le droit pour tout citoyen de pouvoir réutiliser librement et gratuitement les données publiques, ou l'encadrement strict des redevances de réutilisation de données publiques.

En intégrant la mission Etalab au SGMAP, il a constitué une force globale d'innovation, associant la conception des systèmes d'information de l'Etat, l'ouverture des données publiques, l'évaluation des politiques publiques et l'innovation en matière de politiques publiques.

Dès le premier comité interministériel pour la modernisation de l'action publique (CIMAP), le 18 décembre 2012, le gouvernement a pris trois décisions importantes pour l'ouverture des données publiques :

- « Réaffirmer le principe de gratuité de la réutilisation des données publiques et l'étendre, en concertation, aux collectivités, aux délégataires de service public, à la sphère sociale et aux autorités administratives indépendantes » ;
- « Lancer, dès 2013, des travaux sur la transparence de l'action publique et des institutions, notamment en ouvrant le débat sur la mise à disposition de données en matière d'éducation, de risques environnementaux, de transparence du système de santé, de l'offre de transport, de logement, de lisibilité des prestations sociales et des dépenses publiques » ;
- « Confier au SGMAP, en lien avec le ministère du budget, la mission d'évaluer les modèles économiques de chaque redevance existante, notamment en auditant les coûts et les recettes associés. Le SGMAP doit remettre ses propositions au Premier ministre au printemps 2013. »

L'ouverture des données publiques est engagée. Outre les informations partagées (mais pas toujours réutilisables) par le service public de l'accès au droit (Légifrance), le cadastre ou encore Géoportail, de nombreuses bases de données sont mises en libre accès par les administrations, les établissements publics et les collectivités locales.

La mission Etalab a développé une plateforme nationale d'hébergement et de partage de données publiques ouvertes, gratuites et réutilisables (www.data.gouv.fr). Elle coordonne un réseau de

correspondants Open data et de fournisseurs de données qui ont permis la mise en ligne de plus de 350.000 jeux de données. La mission Etalab a également travaillé à structurer l'écosystème des réutilisateurs de données (startups, associations, ...), notamment par le biais de l'organisation du concours Dataconnexions, qui lui a permis, en deux éditions, de récompenser 11 projets particulièrement prometteurs.

L'ambition du Gouvernement en matière d'ouverture des données publiques

Le gouvernement entend donner sa pleine mesure à cette politique, et en faire un levier de modernisation de l'action publique et de soutien au dynamisme économique.

1. Travailler à l'ouverture de jeux de données stratégiques

Le gouvernement souhaite favoriser l'ouverture de données stratégiques et de qualité, qu'elles permettent de créer des services utiles au quotidien (comme pour le transport, l'environnement, la santé, la consommation énergétique, le tourisme...), qu'elles soient réclamées par les citoyens en raison de leur caractère démocratique ou pratique (budgets et dépenses détaillées par administration...), ou qu'elles répondent à une exigence légale ou réglementaire (publication des attributaires de marchés publics conformément à l'article 138 du code des marchés publics).

Pour ce faire, la mission Etalab du SGMAP organisera, dès 2013, six débats thématiques et ouverts, afin d'identifier et de publier de nouveaux jeux de données. Ces débats seront confiés à une commission composée par le SGMAP et le ministère concerné, pouvant bénéficier d'un rapporteur ou d'un secrétariat assuré par Etalab ; ils donneront lieu à un rapport, remis au Premier ministre et au ministre intéressé, qui formuleront des recommandations en termes d'ouverture cohérente des données et préciseront les bénéfices attendus.

Dans ce cadre, le gouvernement étudiera les possibilités :

- d'élargissement de la stratégie d'ouverture des données publiques à tous les établissements publics administratifs, aux délégataires de service public, aux collectivités territoriales et aux autorités administratives indépendantes ;
- d'extension de la politique d'ouverture des données publiques aux missions de service public à caractère industriel et commercial, aux entreprises publiques et à la sphère associative.

Le Premier ministre a par ailleurs demandé à tous les ministres de préciser leur stratégie d'ouverture des données publiques dans les programmes ministériels de modernisation et de simplification (PMMS), à transmettre d'ici la mi-mars.

2. Faciliter et améliorer le processus d'ouverture des données publiques dans les ministères, les établissements publics, et les collectivités territoriales

Etalab ouvrira, dès 2013, une nouvelle version de la plateforme www.data.gouv.fr. Tout en conservant sa vocation de diffuser largement les données brutes, celle-ci deviendra à terme une vitrine grand public de l'ouverture et du partage des données publiques. Elle proposera en particulier des outils techniques utiles pour s'approprier ces données (outils de visualisation de données, de traitement statistique, etc.).

Une attention toute particulière sera apportée à l'indexation des données, à leur traitement sémantique et aux liens entre ces données. Etalab, en collaboration avec les collectivités locales et les autres producteurs de données publiques, travaillera à répertorier et à présenter l'ensemble des

données publiques librement accessibles et gratuitement réutilisables, quelle qu'en soit la source. De même, Etalab poursuivra avec les fournisseurs de données la réflexion sur l'importance des formats ouverts et la possibilité de les faire converger, et s'attachera à construire une plus grande interopérabilité entre la plateforme www.data.gouv.fr et les systèmes d'information de l'État, des collectivités territoriales ou des institutions européennes.

Par ailleurs, Etalab lancera également dès 2013 le chantier de la diffusion de données par des interfaces programmables (« API »).

3. Accompagner les réutilisations innovantes et la communauté des réutilisateurs

L'État encourage la réutilisation la plus large des données publiques afin que cette politique d'ouverture soutienne une stratégie d'innovation, y compris sociale, et de croissance économique.

• Le service public est le premier utilisateur de données publiques de qualité

Les données publiques sont produites dans le cadre d'une mission de service public et le service public est fréquemment le premier bénéficiaire de leur ouverture. La production, la collecte et la publication de données par l'administration doivent répondre aux besoins d'un service public de qualité. Etalab doit donc être à la disposition des administrations pour leur permettre de saisir ces opportunités. En particulier, elle organisera des sessions de formation, pourra monter des ateliers ad-hoc, pour analyser le potentiel de certaines applications de l'ouverture des données publiques (lutte contre le chômage, par exemple), et publiera des notes d'analyses de cas exemplaires.

• L'Etat favorise la réutilisation des données publiques par les entreprises pour créer des services innovants

Etalab continuera à soutenir les entreprises innovantes, l'économie sociale et solidaire ainsi que les grands groupes industriels, dans leurs projets de réutilisation de données publiques. Elle organisera en 2013 trois concours Dataconnexions, permettant de repérer et de faire connaître des réutilisations exemplaires, et soutiendra l'organisation d'Assises nationales de l'open data, en lien avec le ministère chargé du redressement productif, le ministère chargé des PME et le ministère chargé de l'économie sociale et solidaire.

• L'Etat soutient l'engagement des citoyens et des associations autour des données publiques au profit d'une action publique moderne et renouvelée

Le soutien à l'engagement des citoyens, pour coproduire certaines informations, concevoir de nouveaux services ou aider les acteurs publics à relever certains défis, fait partie intégrante des missions d'Etalab, qui organisera en particulier des concours mobilisant les citoyens autour des données et des objectifs partagés par l'État.

• Suivre et encourager la recherche en sciences de la donnée

Les « Big Data », la transformation des chaînes de valeur, les stratégies et les organisations fondées sur la « culture de la donnée » caractérisent les formes actuelles de la révolution numérique. Pour suivre cette transformation rapide, Etalab accentuera sa coopération avec le monde de l'enseignement supérieur et de la recherche et se dotera, avant la fin de l'année 2013, d'un collège d'experts rassemblant des experts technologiques, des spécialistes de datasciences, des juristes, des économistes ou des réutilisateurs (entrepreneurs ou responsables associatifs), à moyens constants.

4. La transparence et la culture de la donnée au service de la modernisation de l'action publique

L'ouverture des données publiques est un outil majeur au service de la transformation du service public, pour en améliorer la qualité et en permettre l'évaluation par tous ses usagers. Le SGMAP veillera à construire, chaque fois que cela sera pertinent, des référentiels quantitatifs, si possible en temps réel, pour les évaluations des politiques publiques, afin qu'elles permettent de construire des outils utiles au pilotage de ces politiques publiques. Comme prévu par le CIMAP du 18 décembre 2012, le SGMAP travaillera à la conception d'innovations dans les services publics en ayant recours aux données publiques.

5. L'évaluation des redevances existantes

À l'occasion du CIMAP du 18 décembre 2012, le Premier ministre a confié au SGMAP, en lien avec le ministère du Budget, une mission d'évaluation des modèles économiques des redevances existantes, notamment en auditant les coûts et les recettes associés. Ces recommandations seront remises au premier semestre 2013.

Etalab et le SGMAP accompagneront les administrations qui commercialisent de l'information publique dans l'évolution de leurs modèles économiques. En effet, le respect du principe d'accès libre et gratuit du citoyen à des données publiques réutilisables n'empêche pas de concevoir des stratégies de monétisation de services à haute valeur ajoutée ou de création de plateformes.

6. Évolutions du cadre juridique et administratif

Un vade-mecum rappelant le cadre juridique de la mise à disposition des données publiques sera diffusé aux ministères, afin de faciliter leur engagement dans la démarche d'ouverture de leurs données.

A l'occasion, notamment, de la révision de la directive 2003/98/CE du Parlement européen et du Conseil du 17 novembre 2003 concernant la réutilisation des informations du secteur public, le gouvernement lancera des travaux de concertation, visant à actualiser si nécessaire les définitions légales encadrant le droit d'accès et de réutilisation des données publiques.

7. L'action internationale de la France

Enfin, le gouvernement attache une importance toute particulière au rayonnement de la France sur le sujet de la transparence, de l'ouverture des données publiques (« Open Data ») et de la gouvernance ouverte (« Open Government »). Etalab prendra part au pilotage des projets européens de plateformes de données ouvertes communautaires (notamment le site ouvert au début de l'année 2013 par la Commission européenne (<http://open-data.europa.eu/open-data/>)).

La France étudie par ailleurs l'opportunité de rejoindre l'*Open Government Partnership* (« OGP »), pour engager une coopération fructueuse avec les autres gouvernements engagés dans cette voie. Elle se rapprochera du partenariat pour examiner les conditions de son adhésion. La France s'engagera par ailleurs largement dans la coopération multilatérale autour des approches de transparence et de gouvernement ouvert.

Annexe 3b – Décisions en CIMAP relatives à l’ouverture des données publiques (extraits des relevés de décisions)

CIMAP du 18 décembre 2012

Une administration ouverte pour encourager la citoyenneté active

Décision n°32. Réaffirmer le principe de gratuité de la réutilisation des données publiques et l’étendre, en concertation, aux collectivités, aux délégataires de service public, à la sphère sociale, aux autorités administratives indépendantes : dès 2013, la transposition du projet de directive européenne sur la réutilisation des données publiques offrira cette occasion.

Décision n°33. Développer la transparence de l’action publique et stimuler une citoyenneté active. Dès 2013, des travaux seront lancés sur la transparence de l’action publique, en ouvrant le débat sur la mise à disposition des données en matière d’éducation, de risques environnementaux, de système de santé, d’offre de transport et de logement et enfin de prestations sociales. L’implication des citoyens, des usagers et des agents dans la conception des politiques publiques et l’élaboration des services publics sera favorisée.

Décision n°34. Favoriser l’innovation collective autour des services publics : dix “services numériques de demain” seront conçus et testés dans une démarche de “laboratoire”.

CIMAP du 2 avril 2013

Poursuivre l’ouverture des données publiques à fort impact économique et démocratique

Conformément aux décisions du CIMAP de décembre, la feuille de route du Gouvernement en matière d’ouverture et de partage des données publiques a été rendue publique au cours du séminaire sur le numérique du 28 février 2013. Elle poursuit trois objectifs principaux : l’ouverture et le partage de nouvelles données à fort impact économique et démocratique, la création d’un “écosystème” cohérent de l’ouverture des données publiques et la mise en place d’une ouverture “par défaut”, intégrée dès la conception des systèmes d’information.

De nombreux nouveaux jeux de données ont d’ores et déjà été partagés sur le site data.gouv.fr :

- dans le domaine de la sécurité, les crimes et délits et les faits constatés par les services de police et de gendarmerie en 2012 ;
- de nombreuses données culturelles, telles que la liste des mobiliers classés monument historique par région ou les aides à la presse ;
- dans le champ de la santé, des données ont été mises à jour, telles que les honoraires des médecins libéraux par spécialité et lieu d’exécution de janvier à octobre 2012 ou les dépenses de santé remboursées par l’assurance maladie par région de janvier à octobre 2012 ;
- dans le secteur de la consommation, les cotations des fruits et légumes par marché et par produit sont réactualisées à un rythme hebdomadaire.

Enfin, la mission d’évaluation des modèles économiques des redevances existantes a été lancée et rendra ses conclusions à la fin du mois de juin.

Au-delà de ces premiers acquis, le Gouvernement se mobilise pour concrétiser cette feuille de route :

- Les premiers débats thématiques ouverts (logement, santé, enseignement supérieur) permettant de créer les conditions d’une politique d’ouverture des données publiques utile aux citoyens, aux entreprises et à la modernisation de l’action publique seront lancés au premier semestre 2013 ;

- Un vade-mecum de l'ouverture des données publiques, rappelant le cadre juridique et technique de la mise à disposition des données publiques, sera mis à disposition des administrations avant juin 2013 ;
- L'étude technique des modalités de l'adhésion au "Partenariat pour une gouvernance ouverte" (Open Government Partnership) est en cours.

Pour franchir une étape supplémentaire, le CIMAP arrête les décisions suivantes :

Décision n° 13 : Le Gouvernement publiera dès 2013 de nouveaux jeux de données sur data.gouv.fr dans des domaines stratégiques de l'action publique.

Ceci concerne notamment : la sécurité, avec la publication des zones de vigilance accrue en matière de sécurité routière ; la santé, avec la mise en ligne de données sur les permanences de soins ambulatoires et la couverture maladie universelle (CMU) et des indicateurs de qualité et de sécurité des soins ; la vie économique, notamment grâce à la publication de données produites par la Banque publique d'investissement (BPI), la direction des entreprises de la Banque de France et l'observatoire des PME ; la recherche, avec les dépenses en recherches et développement, ainsi que le nombre de chercheurs par région, dans les administrations publiques et privées ; dans le champ environnemental, les données relatives à la qualité de l'eau agrégées par l'ONEMA.

Décision n° 14 : Le Gouvernement prendra systématiquement en compte l'exigence de partage des données publiques dans les réformes qu'il présente.

Ainsi, le projet de loi de décentralisation et de réforme de l'action publique prévoit que les collectivités territoriales qui disposent d'un site Internet mettront les données publiques concernant leur territoire à disposition des citoyens ; l'amélioration des réseaux de transport en Île-de-France sera accompagnée de progrès de "l'information-voyageur" par l'ouverture de données de la SNCF et de la RATP ; le projet de loi sur l'enseignement supérieur et la recherche donnera un plus large accès aux données fiscales aux enseignants-chercheurs.

Décision n° 15 : Le Gouvernement améliorera l'accès aux données ouvertes par la création d'un catalogue national permettant de trouver l'ensemble des données publiques librement accessibles; l'élaboration d'une **nouvelle version de la plateforme www.data.gouv.fr** : un travail collaboratif sera conduit pour faire évoluer le site et en faire un lieu de citoyenneté et d'innovation. Le nouveau site sera opérationnel à l'automne 2013.

CIMAP du 17 juillet 2013

L'ouverture des données publiques s'intensifie

Depuis le CIMAP du 2 avril 2013, le Gouvernement a mis à disposition plus de 200 nouvelles séries de données portant sur des domaines stratégiques de l'action publique, comme par exemple : l'aide publique au développement (données détaillées projet par projet), la transparence démocratique (résultats du référendum du 7 avril 2013 relatif au projet de création d'une collectivité territoriale d'Alsace), la transparence de l'action de l'Etat (attributaires de marchés publics), l'égalité entre les femmes et les hommes (principaux chiffres clés en 2012, panorama régional des inégalités), l'éducation (résultats du diplôme national du brevet, classement des performances des lycées), la santé (catégories majeures de diagnostic par établissement de santé), le tourisme (ensemble des établissements touristiques classés) ainsi que la liste de l'ensemble des interventions économiques auprès des entreprises étudiées dans le cadre de la mission d'évaluation de cette politique.

Par ailleurs, sur la base du rapport de M. Mohammed Adnène Trojette, magistrat à la Cour des comptes, relatif au modèle économique des redevances d'utilisation des données publiques, qui sera

remis au Gouvernement et rendu public, le Gouvernement élaborera un plan d'action partagé relatif aux redevances existantes.

Décision n°16 : A la suite de la signature par le Président de la République de la charte du G8 pour l'ouverture des données publiques, le 18 juin 2013, le Premier ministre demande à chacun des ministres de participer à l'élaboration du plan d'action national prévu par la charte pour la fin 2013.

Décision 17 : Le SGMAP (mission Etalab) coordonnera avec l'ensemble des administrations concernées la transposition dans le droit national de la directive européenne du 26 juin 2013 concernant la réutilisation des informations du secteur public.

Décision n°18 : Les trois premiers débats thématiques sur l'accès aux données publiques sont effectivement programmés :

- Le ministère de l'égalité des territoires et du logement (METL) confie au conseil national de l'habitat (CNH), organe consultatif largement ouvert aux parties prenantes du logement, le soin de mettre en place, avec l'appui de la mission Etalab, un groupe de travail pluraliste (composé de membres du CNH et de représentants du monde des réutilisateurs de données publiques et de spécialistes de l'open data) qui commencera ses travaux au mois de septembre 2013 et qui remettra ses préconisations au METL au début de l'année 2014.
- Le ministère des affaires sociales et de la santé lancera un débat thématique sur l'ouverture des données publiques de santé. Nourri des conclusions de la mission confiée à M. Pierre-Louis Bras, ce débat sera officiellement lancé au mois de septembre 2013 pour une remise de conclusions au plus tard dans les six mois. Il sera porté par une commission pluraliste associant les parties prenantes pertinentes et ouvert aux contributions citoyennes.

CIMAP du 18 Décembre 2013

Open data et Gouvernement ouvert : faire de l'administration une plateforme pour l'innovation économique et sociale Le Gouvernement réaffirme son engagement dans le projet d'open data et de Gouvernement ouvert, levier de confiance démocratique, de stimulation de la croissance, d'innovation et de modernisation de l'action publique.

Le Gouvernement ouvre ce jour un nouveau portail d'open data, à l'adresse data.gouv.fr. Ce nouveau portail ouvre la voie d'une coproduction des données d'intérêt public. Il rassemble des données publiques et des données produites par les forces vives de la société, et met en avant les réutilisations et enrichissements de ces données apportés par la société civile (citoyens, associations...), les innovateurs, les chercheurs, les journalistes, etc.

Pour accélérer l'ouverture et le partage des données publiques, le Gouvernement clarifie sa doctrine en matière de redevances sur la réutilisation de données publiques :

Décision n°24 : Le Gouvernement réaffirme le principe de gratuité de la réutilisation des données publiques et décide de ne plus autoriser la création de nouvelle redevance.

Décision n°25 : Sur la base du rapport remis par M. Trojette relatif au modèle économique des redevances de réutilisation des données publiques et après examen des différentes catégories de données publiques soumises à redevance, le Gouvernement décide de supprimer plusieurs redevances. Les redevances supprimées sont celles perçues sur :

- les données relatives aux dotations globales de fonctionnement versées aux collectivités locales par la direction générale des collectivités locales (ministère de l'Intérieur) ;

- les données comptables brutes gérées par la direction générale des finances publiques – balance des comptes définitifs des collectivités et établissements publics locaux recensés à compter de l'arrêté

définitif des comptes 2013. Les fiches de situation financière des collectivités locales restent consultables sur le site collectivites-locales.gouv.fr

- le parcellaire agricole de l'Agence de services et de paiement (ministère de l'agriculture, de l'agroalimentaire et de la forêt) ;
- les données cadastrales de l'Institut national de l'origine et de la qualité (ministère de l'agriculture, de l'agroalimentaire et de la forêt) ;
- les données brutes diffusées par l'Institut français du cheval et de l'équitation ;
- les données relatives aux adresses postales des établissements scolaires (ministère de l'éducation nationale) ;
- les données de l'Office national d'information sur les enseignements et les professions – ONISEP (ministère de l'éducation nationale) ;
- les données de la CADA
- les redevances sur la réutilisation professionnelle des informations suivantes diffusées par la DILA :

ASSOCIATIONS, DEPOTS DES COMPTES DES ASSOCIATIONS, JORF, LEGI, KALI, JADE, CONSTIT, CASS-INCA, CAPP, CNIL, CIRCULAIRES.

Par ailleurs, le ministère de l'économie et des finances fournira en open data, sur une base hebdomadaire, la base des prix des carburants.

Décision n°26 : Le Gouvernement précise sa **doctrine en matière d'exceptions au principe de gratuité** :

- aucune redevance ne saurait être exigée sur des données résultant des missions de service public des administrations générales. Il est demandé aux administrations d'accroître leurs efforts en matière de publication de données leur permettant de mener à bien leurs missions, de motiver leurs décisions, de mesurer leurs activités ou leur bilan. Le site data.gouv.fr, est en mesure d'accueillir ces données ;
- le respect du secret statistique et donc l'anonymisation des informations pouvant éventuellement permettre d'identifier des personnes est un principe central du droit français. Certaines redevances ont été instaurées pour financer cette tâche, qui est indispensable avant ouverture de la donnée. Le Gouvernement demande au SGMAP d'apporter tout le soutien possible aux services publics concernés, afin d'identifier les moyens de réduire ces coûts sans dégrader la qualité de cette anonymisation ;
- enfin, le Gouvernement réaffirme que les opérateurs dont la mission même est de produire des données doivent rechercher des modèles économiques leur permettant de faire face à un paysage économique en profonde reconstitution. Conformément aux conclusions du rapport Trojette, il leur demande d'engager, dans les meilleurs délais, avec l'appui du SGMAP et du ministère du Budget, une réflexion sur les évolutions de leurs modèles économiques. Il leur demande de rechercher des modèles stimulant l'innovation autour de leurs données, favorables aux entrepreneurs innovants, et soutenables à l'heure de l'économie numérique, de la production de nombreuses données par les citoyens eux-mêmes, et des stratégies de plateformes.

Afin de mener cette transition - qui vise à, maintenir et améliorer la qualité des informations publiques, à garantir la pérennité du service public et à développer des stratégies numériques pérennes, efficaces et innovantes, le Gouvernement demande au SGMAP d'accompagner au mieux ces opérateurs, tel l'IGN (Institut géographique national), le SHOM (Service hydrographique et océanographique de la marine) ou Météo France, en lien avec les ministères de tutelle pour définir avec eux des trajectoires d'investissement dans l'innovation compatibles avec le redressement des finances publiques.

Le Gouvernement fait le choix d'investir durablement au service de la transition numérique de l'Etat et s'apprête à lancer les premiers appels à projets du Programme d'investissements d'avenir « Transition numérique et modernisation de l'action publique », doté de 150 M€.

Décision n°27 : Le Gouvernement lancera les premiers appels à projets co-financés par le **Programme d'investissements d'avenir « Transition numérique et modernisation de l'action publique »**, au premier semestre 2014. Les projets seront évalués par un comité d'experts indépendants et sélectionnés dans le cadre d'une gouvernance interministérielle mise en place avec le Commissariat général à l'Investissement et le SGMAP. Ils feront l'objet d'un suivi technique et financier régulier et donneront lieu à une évaluation ex post.

Annexe 4
La conservation, la transmission et l'hébergement des données de santé

Tout responsable de traitement de données à caractère personnel est soumis à une obligation générale de sécurité, définie par l'article 34 de la loi du 6 janvier 1978, selon laquelle « *le responsable du traitement est tenu de prendre toutes précautions utiles, au regard de la nature des données et des risques présentés par le traitement, pour préserver la sécurité des données et, notamment, empêcher qu'elles soient déformées, endommagées, ou que des tiers non autorisés y aient accès* ». Toutefois, compte tenu de la sensibilité particulière des données de santé, le législateur a instauré, dans le cadre de la loi du 4 mars 2002 relative aux droits des malades, deux régimes spécifiques destinés à préciser et à renforcer les obligations pesant sur les personnes assurant le traitement de ces données :

- le premier dispositif, prévu par l'article L. 1110-4 du code de la santé publique, régit la conservation sur support informatique et la transmission par voie électronique des données médicales détenues par les professionnels et les établissements de santé ;
- le second, qui figure à l'article L. 1111-8 du même code, encadre le dépôt de données de santé par les professionnels et les établissements de santé auprès de personnes tierces, qualifiées d'hébergeurs.

S'agissant de la conservation et de la transmission des données médicales, l'article R. 1110-1 dispose qu'elles doivent s'effectuer dans le respect de « *référentiels* » pris par arrêté du ministre de la santé après avis de la CNIL, qui fixent le niveau de sécurité requis. L'utilisation de la carte de professionnel de santé ou (depuis la loi HPST) d'un dispositif équivalent est obligatoire pour accéder aux données médicales informatisées et pour les transmettre par voie électronique.

S'agissant de l'hébergement de données de santé, la loi réserve cette activité à des prestataires agréés par le ministre de la santé, après avis de la CNIL et d'un comité d'agrément¹. L'agrément permet notamment de s'assurer de la qualification des personnels, de la sécurité des solutions techniques et de la mise en œuvre d'une politique de confidentialité. En outre, l'hébergement de données de santé ne peut se faire qu'avec le consentement exprès de la personne concernée. Les établissements et les professionnels de santé ne peuvent déroger à cette exigence de consentement que si l'accès aux données est limité à l'établissement ou au professionnel qui les a déposées. 64 sociétés ont été agréées à ce jour².

On peut souligner que le régime de l'hébergement de données de santé est une « *règle relative aux services* »³ au sens de la directive n° 98/34/CE du Parlement européen et du Conseil du 22 juin 1998 prévoyant une procédure d'information dans le domaine des normes et réglementations techniques et des règles relatives aux services de la société de l'information. En vertu des articles 8 et 9 de la directive, tout projet de modification significative de ce régime⁴ doit donc être communiqué à la Commission européenne et son adoption ne peut intervenir moins de trois mois avant la réception du projet par la Commission.

¹ Ce comité comprend un membre de l'IGAS, deux représentants des associations de patients, deux représentants des professionnels de santé et trois personnalités qualifiées compétentes en matière d'éthique et de droit, de sécurité des systèmes d'information et dans le domaine économique et financier.

² ASIP santé, mars 2014.

³ Article 1.5 : une règle relative aux services est « *une exigence de nature générale relative à l'accès aux activités de services visées au point 2 et à leur exercice, notamment les dispositions relatives au prestataire de services, aux services et au destinataire de services, à l'exclusion des règles qui ne visent pas spécifiquement les services définis au même point* ».

⁴ Selon l'article 8.1, la communication d'une modification d'une règle technique s'impose si les Etats « *apportent au projet de règle technique, d'une manière significative, des changements qui auront pour effet de modifier le champ d'application,*

d'en raccourcir le calendrier d'application initialement prévu, d'ajouter des spécifications ou des exigences ou de rendre celles-ci plus strictes ».

Annexe 4 bis
Les dossiers de santé personnalisés

Un dossier de santé personnalisé rassemble sur un même traitement informatique des données à caractère personnel concernant un patient issues de l'activité de différents professionnels de santé. Trois systèmes coexistent aujourd'hui, qui ont chacun leur cadre légal et réglementaire : le dossier médical personnel (DMP) ; le dossier pharmaceutique (DP) ; le service « Web Médecin ».

1) Le dossier médical personnel

La création du dossier médical personnel a été prévue par la loi du 13 août 2004 relative à l'assurance-maladie. Il a vocation à recevoir l'ensemble des informations à caractère médical concernant à une personne (recensement des actes et consultations effectués par chaque professionnel de santé, éléments diagnostiques et thérapeutiques, résumés de séjour en établissement de santé) et à les conserver tout au long de sa vie.

Le régime du DMP est marqué par le souci de respecter le secret médical et l'exigence de consentement de la personne concernée. La création du DMP, qui peut être effectuée par tout professionnel de santé, nécessite l'accord de la personne. Celle-ci doit en outre donner son autorisation d'accès pour chaque professionnel de santé. L'accès au DMP ne peut être donné, même avec l'accord de la personne, à des personnes autres que les professionnels de santé exerçant en ville ou en établissement. Il est notamment interdit lors de la conclusion d'un contrat relatif à une complémentaire santé ou nécessitant l'évaluation de l'état de santé de la personne, ainsi que dans le cadre de la médecine du travail. Il ne peut être dérogé à l'exigence de consentement à l'accès que si la personne est hors d'état d'exprimer sa volonté et en présence d'une situation comportant un risque immédiat pour sa santé ; la personne peut s'opposer de manière expresse à ce que les professionnels de santé accèdent au DMP dans cette hypothèse.

La CNIL a autorisé une première phase de déploiement généralisé du DMP par une délibération n° 2010-449 du 2 décembre 2010. Plusieurs garanties pour la personne concernée ont été précisées dans le cadre du dossier de demande d'autorisation et de la décision de la CNIL : exigence d'une copie papier du document par lequel le professionnel de santé ayant ouvert le DMP atteste avoir informé la personne et recueilli son consentement ; remise préalable à la personne d'une brochure d'information, précisant notamment que son refus de créer un DMP n'aura pas de conséquences sur ses remboursements ainsi que les possibilités d'accès de la personne au DMP depuis son ordinateur personnel ; droit de la personne de demander que certains documents ne soient pas mentionnés dans le DMP ou soient « masqués » à certains professionnels de santé.

Le DMP est hébergé par un hébergeur agréé de données de santé (cf. annexe 1). Suite à un appel d'offres lancé en 2009 par l'ASIP santé, maître d'ouvrage du DMP, un groupement composé de la société Santeos et de filiales d'Atos et de La Poste a été retenu.

Plusieurs éléments du cadre juridique du DMP restent en suspens. Ni le décret prévu par l'article L. 1111-21 du code de la santé publique, qui devait notamment fixer les conditions d'accès aux différentes catégories d'information, ni celui relatif à l'identifiant national de santé, numéro distinct du NIR et non signifiant et qui devait être utilisé pour le DMP et le DP, n'ont été adoptés. Dans sa délibération du 2 décembre 2010, la CNIL a admis que le cadre juridique ne soit pas complètement fixé dans l'attente des retours d'expérience, mais a estimé que la deuxième phase de déploiement du DMP devrait se faire dans un cadre juridique achevé. Des dispositions législatives modifiant ponctuellement le cadre actuel seront inscrites dans la loi de santé publique et un décret d'application sera publié en aval.

A ce jour, près de 450 000 DMP ont été créés⁵.

2) Le dossier pharmaceutique

La création d'un dossier pharmaceutique a été prévue par une loi du 30 janvier 2007. L'objet du dossier pharmaceutique est plus restreint que le DMP : il ne comporte que les informations relatives aux médicaments et dispositifs médicaux dispensés. Si son ouverture est subordonnée, comme pour le DMP, au consentement de la personne attesté par la remise d'un écrit, son utilisation est ensuite soumise à des conditions moins strictes : alors que pour le DMP, le patient doit donner son accord à l'accès de chaque professionnel de santé (système qu'on peut qualifier « *d'opt-in* »), pour le DP, tout pharmacien est en principe tenu de le consulter et de l'alimenter, sauf opposition de la personne (système « *d'opt-out* »).

Le DP permet notamment aux pharmaciens de prendre connaissance des traitements pris par la personne et déceler des risques d'interactions médicamenteuses indésirables. Dans le cadre d'une convention signée le 3 novembre 2011 entre le Conseil national de l'ordre des pharmaciens, maître d'ouvrage du dispositif, et l'ANSM, il sert également à la diffusion des messages de rappels de lots de médicaments (fonction « DP-Rappels »). Enfin, l'article 23 de la loi du 29 décembre 2011 relative au renforcement de la sécurité sanitaire du médicament et des produits de santé a autorisé le ministre de la santé, l'ANSM et l'INVS à accéder aux données anonymes contenues dans les DP, dans une optique de veille sanitaire.

Le décret n° 2012-1131 du 5 octobre 2012 a étendu le DP aux pharmaciens exerçant dans les pharmacies à usage intérieur des établissements de santé. En outre, une expérimentation relative à l'accès au DP de médecins exerçant dans certains établissements de santé a été lancée pour trois ans par le décret n° 2013-31 du 9 janvier 2013 ; elle concerne notamment les médecins anesthésistes-réanimateurs, les urgentistes et les médecins exerçant dans les structures de médecine gériatrique. En revanche, l'alimentation du DMP par le DP, prévue par la loi, est subordonnée à l'adoption du décret relatif aux conditions de mise en œuvre du DMP (cf. point précédent).

Fin 2012, 97 % des officines de pharmacie étaient raccordées au DP et 24 145 000 DP avaient été créés, soit un taux de couverture de 36 %. Le taux de couverture est supérieur à 50 % pour les plus de 60 ans⁶.

3) Le service « Web Médecin »

Comme le DMP, le service Web Médecin a été créé par la loi du 13 août 2004 relative à l'assurance-maladie. Contrairement au DMP et au DP, il s'agit d'un service exclusivement destiné aux médecins, qui n'est pas directement accessible aux patients⁷. Il permet aux médecins de consulter l'historique de remboursements de leurs patients sur les douze mois précédant la consultation. Le patient doit donner son consentement à cette consultation, consentement qui est réputé obtenu lorsque le médecin utilise à cette fin la carte Vitale.

La gestion technique du dispositif est assurée par la CNAMTS.

⁵ 445 157 le 26 mars 2014, selon le site www.dmp.gouv.fr

⁶ CNOP, *Le dossier pharmaceutique. Rapport d'activité de l'année 2012*, avril 2013.

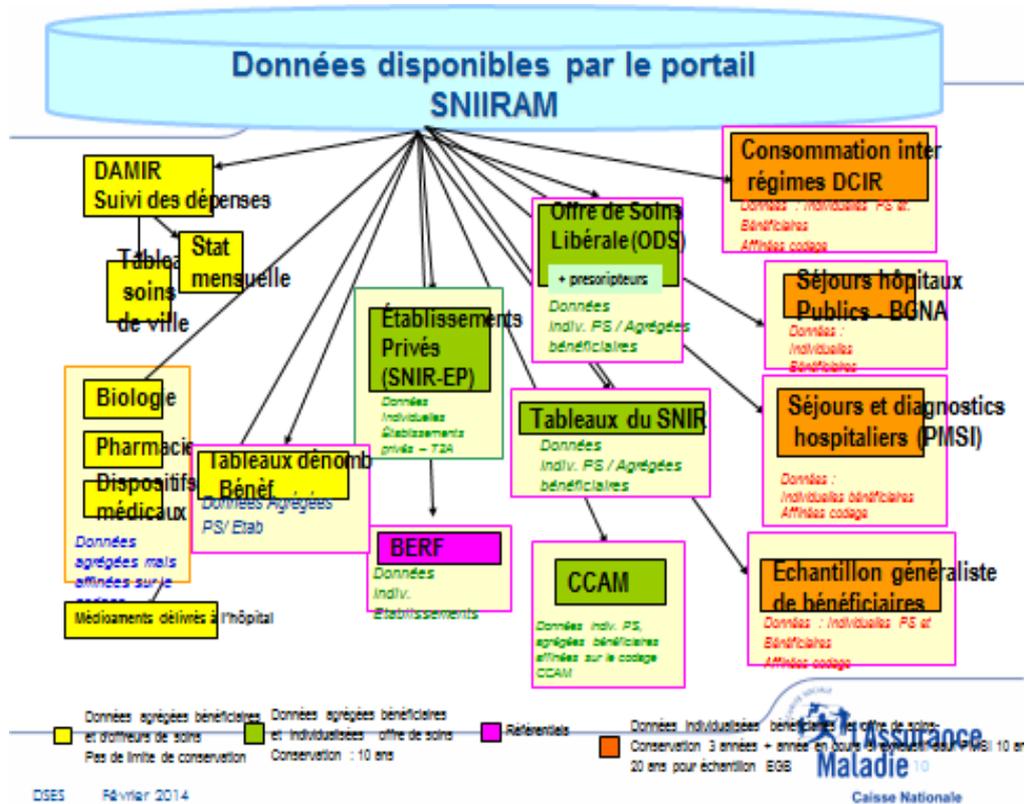
⁷ Le patient dispose cependant, comme pour tout traitement de données à caractère personnel, d'un droit d'accès qui s'exerce par demande adressée à la caisse d'assurance-maladie dont il relève.

Annexe 5

Arrêtés SNIIRAM du 19 juillet 2013 et du 14 février 2014

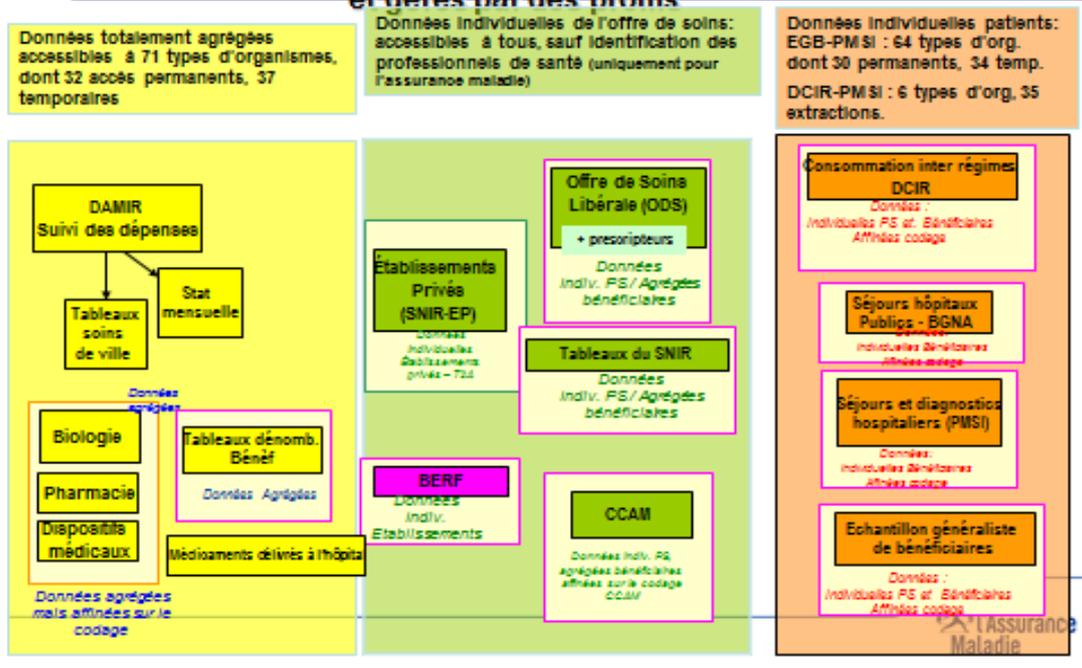
	Suivi des dépenses	Suivi de l'activité offreur soins	de par de	Echantillon généraliste des bénéficiaires	Données individuelles exhaustives
Caisses des régimes de base d'assurance maladie + CNSA	X	X		X	X
Institut de veille sanitaire (InVS)	X	X		X	X
Agence nationale de sécurité du médicament (ANSM)	X	X		X	X
Haute autorité de santé (HAS)	X	X		X	X
Agences Régionales de Santé (ARS)	X	X		X	X (sur périmètre régional)
Ministères (santé, sécurité sociale et finances) et services déconcentrés	X	X		X	(en cours DREES)
Haut conseil pour l'avenir de l'assurance maladie (HCAAM)	X	X		X	
Agence de la biomédecine	X	X		X	
Institut national du cancer (INCa)	X	X		X	
Agence technique de l'information sur l'hospitalisation (ATIH)	X	X		X	(en cours ATIH)
Union nationale des professions de santé (UNPS)	X	X		X	
Institut national de la santé et de la recherche médicale (INSERM)	X	X		X	
Centre national de la recherche scientifique (CNRS)	X	X		X	
Institut de Recherche et Documentation en Economie de la Santé (IRDES)	X	X		X	
Centre technique d'appui et de formation des centres d'examens de santé (CETAF)	X	X		X	
Fonds CMU	X	X		X	
Observatoire des drogues et toxicomanies	X	X		X	
Institut des données de santé (IDS)	X	X		X	
Union Nationale des Organismes Complémentaires d'Assurance Maladie (UNOCAM) <i>Membre de l'IDS</i>	X	X		X	
Fédérations constitutives de l'UNOCAM (FNMF, CTIP, FFSA) signataires de la charte d'engagement entre l'assurance maladie obligatoire et les assureurs maladie complémentaires	X	X		X	
Fédérations hospitalières (FHF, FHP, FEHAP) <i>Membre de l'IDS</i>	X	X		X	
Collectif interassociatif sur la santé CISS <i>Membre de l'IDS</i>	X	X		X	
Syndicats de professionnels libéraux membres de l'UNPS	X	X			
Unions régionales des professions de santé (URPS)	X	X			
Etablissements membres des fédérations hospitalières	X	X			
Associations adhérentes au CISS	X	X			
Organismes complémentaires contributeurs de données membres des fédérations constitutives de l'UNOCAM signataires de la charte d'engagement entre l'assurance maladie obligatoire et les assureurs maladie complémentaires	X	X			

Annexe 5 bis
Cartographie des produits SNIIRAM



Cartographie des types d'accès

Caractéristiques des accès SNIIR-AM définis par un arrêté et gérés par des profils



Annexe 6

Cartographie des données de santé

L'étude réalisée par le SGMAP est disponible à l'adresse :

<https://www.data.gouv.fr/fr/dataset/cartographie-des-bases-de-donnees-publiques-en-sante>

Annexe 7

Panorama international et principaux enseignements

On trouvera en document joint une synthèse de l'étude :



[Panorama de l'open data en santé – note de synthèse.pdf](#)

La présentation détaillée du panorama se trouve à l'adresse :



[Open data en sante-panorama international 12 mars 2014.pdf](#)

Annexe 8

La cartographie des données de santé

Annexe 7a – Liste des institutions rencontrées pour réaliser la cartographie des données de santé

La liste des institutions rencontrées est exhaustive. Toutes les personnes rencontrées au sein de chaque institutions n'ont cependant pas pu être mentionnées ici. Le SGMAP remercie l'ensemble des personnes rencontrées pour le temps qu'ils nous ont consacré.

Institution	Interlocuteur	Position	Date
DREES	Franck Von Lennep	Directeur	05-févr
DREES	André Loth	Directeur de projet	05-févr
DREES	M. Badouraly	Responsable SI	05-févr
IDS	Richard Decottignies	Directeur adjoint	05-févr
INSERM	Grégory Rey	Directeur du CépiDC	06-févr
ATIH	Max Bensadon	Directeur adjoint	07-févr
INPES	Pierre Ardvison	Directeur des affaires scientifiques	07-févr
IRDES	Yann Bourgeuil	Directeur	10-févr
INSERM	Marcel Goldberg	Chercheur / expert	11-févr
HAS	Catherine Grenier	Chef du service indicateurs pour l'amélioration de la qualité et de la sécurité des soins	17-févr
HAS	Sabine Cohen-Hygounenc	Responsable de la mission d'information sur la qualité des prises en charge dans les établissements de santé à destination des usagers	17-févr
HAS	Etienne Doutreleau	Adjoint au secrétaire général	17-févr
HAS	Georges Dutreix	Chef de projet SSI	17-févr
HAS	Hervé Nabarette	Chef du service Qualité de l'information médicale	17-févr
HAS	Ariane Sachs	Service évaluation et amélioration des pratiques	17-févr
ASIP Santé	Jeanne Bossi	Secrétaire Générale	17-févr
ASIP Santé	Anne Monnier		17-févr
ANSM	Mahmoud Zureik	Directeur de la stratégie et des relations internationales	18 février et 28 mars
CGIET	Robert Picart	Référent Santé au CGIET	19-févr
	Fabrice Mattatia		19-févr
INVS	Anne Bruant-Bisson	Directrice générale adjointe	21-févr
INVS	Javier Nicolau	Chef du projet informatique	21-févr
INVS	Anne Doussin	Directrice du département des maladies chroniques	21-févr
INSEE	Thierry Aouizerate	Adjoint chef du dpt info service	24-févr
INSEE	Pascale Breuil	Chef de l'unité des Etudes démographiques et sociales	24-févr
INCA	Philippe-Jean Bousquet	Responsable du département Veille, Observation et Documentation	24-févr
INCA	Fabien Calvo	Directeur général adjoint, directeur de la recherche	24-févr
CNSA	Juliette Bloch	Directrice scientifique	25-févr
CNSA	Jean-Luc Bernard	Directeur adjoint SI	25-févr
CNSA	Hamid Boukhrouf	Directeur des SI	25-févr
DGS	Marie-Christine Favrot	Adjoint au directeur général, chef des politiques de santé	25-févr
DGS	Anne Périllat	Adjointe au chef du bureau du médicament	25-févr
DGS	Alain Fontaine		25-févr

Institution	Interlocuteur	Position	Date
ARSIF	Patrick Olivier	Directeur de la stratégie	10-mars
ARSIF	David Clavereau	Directeur de projet e-santé	10-mars
FHF	Yves Gaubert	Responsable pôle finances	28-févr
DSSIS	Philippe Burnel	Délégué à la stratégie des systèmes d'information en santé	05-mars
CNAMTS	Claude Gissot	Directeur de la stratégie, des études et des statistiques	11-mars
CNAMTS	Hélène Caillol		11-mars
ANAP	Christian Anastasy	Directeur	11-mars
ANAP	Gilles Bontemps	Adjoint directeur	11-mars
DSS	François Godineau	Adjoint au directeur	05-mars
DSS	Abraham Hamzaoui	Chef de bureau des SI	05-mars
Malakoff Mederic	Laurent Borella	Directeur santé	19-mars
Malakoff Mederic	Laurence D'haultfoeuille		19-mars
DGOS	Gaël Raimbault	adjoint sous directeur pilotage	26-mars
DGOS	Yannick Leguen	sous directeur pilotage	26-mars
FHP	David Castillo	Direction de la prospective économique, médicale	20-mars
FNORS	Alain Trugeon	ORS Picardie	31-mars
ARS PACA	Norbert Nabet	Directeur Général adjoint	08-avr
Réseau Quêtelet	Roxanne Silberman	Secrétaire Générale	14-avr
DGOS	Franck Jolivaldt	Mission SI des acteurs de l'offre de soins	26-mars
ARSIF	Laurent Trevillier	Pôle SI	10 mars et 18 avril
ARSIF	David Clavereau	Chef de projet e-santé	10 mars et 18 avril

Les compte-rendus des ateliers se trouvent dans les documents joints :



[Annexe 8b : Compte-rendu des ateliers thématiques – Hôpital.pptx](#)



[Annexe 8c : Compte-rendu des ateliers thématiques - Consommation de soins.pptx](#)



[Annexe 8d : Compte-rendu des ateliers thématiques - Offre de soins.pptx](#)

Annexe 8 bis

Synthèse de la consultation publique



[Synthèse de la consultation publique.pdf](#)

Annexe 9

Rapport du groupe de travail « risques de réidentification »

Conclusions du groupe de travail sur le « Risque de réidentification » dans les bases de données médico-administratives (pour la Commission open data)

A la suite du rapport Bras⁸ qui lui a été remis le 3/10/2013, la ministre, Madame Marisol Touraine, a demandé à Franck Von Lennep, directeur de la recherche, des études, de l'évaluation et des statistiques (DREES), de diligenter une expertise technique sur la sécurité des données concernant le risque de ré-identification des personnes à partir de données [supposées] anonymes.

L'étude a été réalisée de novembre 2013 à avril 2014 par un groupe de travail animé par André Loth (DREES) et composé de : Alireza Banaei et Max Bensadon (ATIH), Hélène Caillol (CNAMTS), Nora et Frédéric Cuppens (Institut Telecom Bretagne), Emmanuelle Denis (DSS), Françoise Dupont (INSEE et CASD), Noémie Jess (DREES), Jean-Pierre Le Gléau (expert), Grégoire Rey (INSERM-CépiDc).

Y ont également contribué (pour le test de jeux de données anonymes) Maxime Bergeat (INSEE), Dominique Blum (expert) et les équipes techniques du CASD.

Le comité de pilotage, présidé par Franck von Lennep, s'est réuni 3 fois. Ont participé à tout ou partie de ces réunions Max Bensadon (ATIH), Philippe Burnel (DSSIS), Anne Coat (ANSSI), Claude Gissot (CNAMTS), Nora et Frédéric Cuppens (Institut Télécom Bretagne), François Godineau (DSS), Philippe Cuneo et Michel Isnard (INSEE), Jean-Pierre Le Gléau (expert), André Loth (DREES), Christian Saout (CISS) ainsi que Marc-André Beaudet (expert de la CNIL, en observateur).

1/ L'objet de l'étude

En substance la mission confiée au groupe de travail était

- d'identifier et évaluer les risques de ré-identification dans les bases médico-administratives considérées (SNIIRAM, PMSI...) notamment pour les jeux de données issus de ces bases et ayant donné lieu à une diffusion relativement large (données hospitalières du PMSI, échantillon généraliste des bénéficiaires) ou dont l'anonymat faisait débat (datamart des consommations inter-régimes de l'assurance maladie) ;
- d'établir une ligne de démarcation entre
 - o jeux de données anonymes (pouvant donc être mis en open data ou publication sans restriction)
 - o jeux de données présentant un risque de ré-identification et devant pour cette raison être en accès restreint...
- De recommander des moyens pour élargir, dans cette offre de données, la part et le volume des données en accès libre ;
- De préciser pour les données en accès restreint (parce que comportant des risques de ré-identification) sous quelles conditions techniques les personnes habilitées à accéder à ces données devraient pouvoir le faire afin de limiter le risque.

⁸ Rapport sur la gouvernance et l'utilisation des données de santé (sur les risques de ré-identification dans les bases de données médico-administratives, voir notamment pp. 26 à 30 et 56 à 58 du rapport).

Le groupe a entendu des chercheurs et parcouru la littérature sur ce sujet. Il a été surpris et déçu de constater qu'il n'existe pas vraiment encore de doctrine établie et que la plupart des responsables de bases de données, publiques et privées mais aussi les responsables des autorités de contrôle (les équivalents de la CNIL) en sont encore à tâtonner. Le groupe des 29 CNIL européennes vient ainsi de définir une liste de trois conditions suffisantes pour qu'un jeu de données puisse être qualifié d'anonyme mais ces trois conditions ne sont pratiquement jamais réunies et on est renvoyé dès lors à une évaluation au cas par cas⁹ : la veille technologique et scientifique sur ces questions devra être poursuivie demain, à l'initiative des gestionnaires des bases mais aussi dans les instances de gouvernance et dans les instances de contrôle afin que la doctrine s'affirme (et évolue quand c'est nécessaire).

2/ L'évaluation du risque de ré-identification dans le cas des données médico-administratives de santé

Il est entendu qu'il n'est question ici que des bases de données dé-identifiées ou « pseudonymisées », dites parfois, à tort, anonymes, ou de « jeux » de données issus de ces bases.

Le risque de ré-identification des personnes dans ces bases ne vient pas tant du pseudonyme utilisé¹⁰ que de la possibilité pour des tiers de reconnaître, dans des jeux de données, des personnes sur lesquelles ils disposent d'informations par ailleurs.

Classiquement (cf. le Modèle d'évaluation harmonisée du risque –MEHARI- du Club de la sécurité de l'information français), le risque, dépend à la fois :

- de la probabilité ou potentialité que l'événement se réalise, malgré des mesures dissuasives ou préventives (potentialité généralement évaluée sur une échelle de 1, faible, à 4, très fort) ;
- de l'impact (noté de 1 à 4 également) qu'a sa réalisation, compte tenu d'éventuelles mesures compensatoires ou réparatrices...

Risque = Potentialité x Impact

S'agissant de l'**impact** d'une divulgation non souhaitée de données personnelles de santé, le groupe a considéré qu'il était *très fort*, s'agissant de données classées comme sensibles dans le droit français et européen, et protégées par le droit constitutionnel à la vie privée. Des exemples de conséquences potentiellement très dommageables pour la personne concernée peuvent être cités (portant sur la carrière professionnelle ou l'image ou la vie de famille notamment).

Le fait que certaines personnes puissent rendre publiques elles-mêmes des informations sur leur propre santé ne change rien à l'évaluation de l'impact pour les victimes d'une divulgation non souhaitée. En outre, une fois l'information divulguée, le mal est fait et on ne peut ni remettre l'information dans la boîte ni considérer qu'une compensation financière réparera le préjudice.

Si on s'accorde à évaluer l'impact comme très fort, cela conduit à refuser en l'espèce une conception utilitariste de l'analyse bénéfice/risque où on admettrait que l'utilité économique pour le plus grand nombre compense le préjudice subi par quelques uns (supposés peu nombreux) dont la vie privée

⁹ <http://www.cnil.fr/linstitution/actualite/article/article/le-g-29-publie-un-avis-sur-les-techniques-danonymisation/>

¹⁰ Dans les bases SNIIRAM et PMSI, le pseudonyme est le NIR chiffré deux fois de manière irréversible. Une bonne organisation (par exemple des clés de chiffrement secrètes ou une table de correspondance, confiées à plusieurs personnes différentes extérieures à la base de données) évite le risque que le pseudonyme permette de remonter à la vraie identité de la personne, identité qui est parfaitement inutile pour les traitements ordinaires. Dans les cas où remonter à l'identité des personnes est nécessaire (pour rendre possible des appariements avec d'autres jeux de données, d'enquête par exemple), le principe de cloisonnement évite que les gestionnaires de la base connaissent les vraies identités. Ce principe de gestion séparée des données directement identifiantes (nom, NIR, numéro de téléphone...) est celui qui est posé par le projet de règlement européen sur la protection des données personnelles (art. 81 et 83).

serait divulguée. Les membres du groupe ont préféré une autre conception de l'analyse bénéfice-risques, consistant à maximiser le niveau de bénéfice pour un niveau de risque maîtrisé. Cela permet de conclure à l'intérêt :

- de rendre les données personnelles de santé accessibles – notamment aux chercheurs ou à des bureaux d'études – dans des conditions convenables (rapidité, assistance...) – mais avec l'accord de la CNIL et en contrôlant le risque de divulgation ;
- d'ajuster le niveau des précautions au niveau du risque ;
- et de maximiser la production et la diffusion des données anonymes issues -par divers procédés- des données individuelles brutes.

S'agissant de la **probabilité de ré-identification** et de l'efficacité de la dissuasion ou des mesures préventives, le groupe a considéré sans s'y appesantir que la diffusion de données agrégées issues des bases médico-administratives ne présentait pas en général de risque de ré-identification et qu'il n'y avait donc pas lieu d'en limiter l'accès à telle ou telle catégorie de personnes ou d'organismes comme c'est encore actuellement le cas, et qu'il y avait en outre certainement des moyens d'accroître la richesse et la pertinence des statistiques mises en ligne.

Comme on le verra, le groupe a également travaillé sur la possibilité d'étendre la production et l'usage de jeux de données individuelles -jeux exhaustifs ou échantillons- ayant pour caractéristiques que le risque de ré-identification y soit a priori inexistant et que ces données puissent donc elles aussi être publiées et réutilisées par tous.

Cependant, il lui était d'abord demandé de se prononcer sur la politique de diffusion des données médico-administratives individuelles par le ministère de la santé et les agences ou établissements publics placés sous sa tutelle. S'agissant des données individuelles à vocation exhaustives du SNIIRAM, du PMSI et de l'ensemble SNIIRAM-PMSI, il confirme l'évaluation concluant à un niveau élevé du risque :

- le risque que présente une diffusion mal contrôlée des données hospitalières du PMSI exhaustif (dans les conditions et avec les précautions insuffisantes¹¹ qui caractérisent son mode de diffusion actuel) ;
- le risque que cela entraîne pour l'Échantillon généraliste des bénéficiaires (EGB) au 1/97, accessible à un nombre important d'organismes, où on trouve les parcours de soins de 600 000 personnes sur (potentiellement) 20 ans. Cet échantillon a été malheureusement compromis par la diffusion trop large du PMSI lequel permet comme on l'a vu d'identifier des personnes hospitalisées et de les retrouver, avec les mêmes caractéristiques dans l'EGB ;
- le risque que présenterait un accès trop large au datamart de consommations inter-régimes (DCIR) exhaustif de l'assurance maladie, si par exemple il devenait ouvert à toutes les personnes ayant le statut de chercheurs, même sans possibilité de croisement des variables sensibles et même sans croisement avec le PMSI (ce qui lui retirerait d'ailleurs une bonne part de son intérêt).

Dans tous ces cas, le nombre de personnes présentant les mêmes caractéristiques au regard des quasi-identifiants (date en mois des soins, durée d'hospitalisation, arrêt de travail, âge en années...)

¹¹ Les calculs de D. Blum, confirmés par l'ATIH, donnent les ratios de 89 % et 100 % de séjours où une personne est seule à présenter les caractéristiques cherchées (respectivement sur l'ensemble des séjours de 2008 et sur les séjours des patients hospitalisés au moins 2 fois dans l'année). Ces calculs ont été effectués en 2011 sur une base présentant le même type de « floutage » que les bases diffusées au cours des années récentes (imprécision volontaire sur la date de sortie et code géographique correspondant à au moins 1000 habitants). La diffusion à plusieurs centaines d'exemplaires de copies de cette base complète, sous la forme de disques, la diffusion systématique du fichier de chaînage des séjours, l'existence avérée de copies sauvages (et l'intérêt d'en obtenir en raison de leur valeur marchande et pour éviter des procédures et un paiement), l'impossibilité pratique d'empêcher ces copies ou de contrôler le respect des engagements pris par les titulaires des autorisations de la CNIL ont été les principaux arguments cités.

est très petit, tendant rapidement vers la valeur 1 si les données sont chaînées dans le temps : **tous les parcours de soins sont différents.**

Si on prend l'exemple de l'INSEE, celui-ci ne prend pas le risque de diffuser un fichier exhaustif présentant ce type de risques autrement que via un dispositif sécurisé à des personnes dont le besoin de connaître les données a été préalablement vérifié et évalué (c'est également ce qui est prévu pour les données fiscales).

Dans cette évaluation de la probabilité de ré-identification, il convient de distinguer trois temps :

- ✓ Le *comptage* du nombre de « personnes dans une case » (personnes présentant les mêmes caractéristiques au regard des quasi identifiants que sont l'âge, le sexe, le lieu de résidence, le lieu, la nature et les dates de soins, la date de décès éventuellement). Il présente un caractère objectif. Ce nombre est fortement réduit jusqu'à devenir vite égal à un pour toute la population dès que les données relatives à une même personne sont chaînées¹² ;
- ✓ Le *classement* des quasi-identifiants par degrés de notoriété n'est guère contestable non plus. Les dates de naissance et de décès par exemple sont aisément connues de tous ; les dates et lieu d'hospitalisation ou les dates d'arrêt de travail restent assez faciles à connaître, même si pour une personne donnée, le nombre de personnes susceptibles de les connaître est généralement petit (proches, employeur, assureur). En revanche, les dates de consultations médicales ou de soins par des auxiliaires médicaux sont plutôt mal connues et vite oubliées ;
- ✓ L'*estimation* de la probabilité que ces informations donnent lieu à une divulgation par les personnes qui en ont eu connaissance est plus subjective (« à dire d'experts ») :
 - cette probabilité augmente évidemment avec le nombre de personnes mises dans le secret, directement ou indirectement, et dépend donc de la procédure de sélection des demandes d'accès ;
 - elle est réduite avec l'emploi de dispositifs techniques permettant le confinement des données, la traçabilité des accès et des requêtes et l'interdiction des copies, alors qu'au contraire la diffusion des données au moyen de copies et d'extractions augmente la probabilité de divulgation par négligence et/ou malveillance ;
 - elle est réduite aussi avec le caractère dissuasif des sanctions pénales pour divulgation illicite, que le groupe a jugé
 - assez efficace dans le cas d'institutions telles que les assurances (secret de pratiques illicites difficile à garder, souci d'image...)
 - mais peu efficace devant d'autres « attaquants » (par exemple adversaires politiques, malveillance dans un cadre familial ou professionnel, presse à scandale, démonstration qu'on peut le faire ou simple curiosité...) surtout si la probabilité d'être pris est très faible (pas de traçabilité pour les données extraites). On notera que les cas de divulgation des données peuvent alors rester confinés à un petit cercle.

En tout état de cause, l'argument « jusqu'ici tout va bien » ne dispense pas, selon le groupe de travail, de se conformer aux standards professionnels en vigueur et ne peut justifier la continuation du mode actuel de diffusion des données du PMSI et de l'EGB, ni l'ouverture incontrôlée des accès au DCIR, qui sont porteuses de risques pour les personnes concernées et qui seraient de nature, en cas d'incidents publics, à saper la confiance dans le système d'information.

Il appartient à la CNIL de faire savoir quelle sera désormais son attitude concernant la diffusion des données du PMSI, mais le groupe de travail recommande au ministère de la santé de revoir sa politique de diffusion des données qui est jugée trop restrictive dans le cas des données agrégées que chacun s'accorde à classer comme anonymes et trop laxiste dans le cas des données du PMSI.

Pour mémoire, le groupe de travail ne s'est pas penché sur le cas des usages opérationnels des données (par exemple dans les caisses d'assurance maladie pour les feuilles de soins ou dans les

¹² Même sans chaînage des données, certains cas facilitent la réidentification des personnes (patients soignés loin de chez eux, décès...).

ARS pour les données du PMSI) : ces usages sont ceux qui ont justifié initialement le recueil des données et sont donc par hypothèse nécessaires au service public. Le groupe rappelle seulement que dans ces cas aussi l'accès aux données doit être encadré (personnes habilitées individuellement et tenues au secret professionnel¹³, tenue de registres, accès sécurisé...).

3) Une réponse graduée selon le risque de ré-identification

L'offre que le groupe de travail propose de faire à ce stade (travaux à poursuivre, en liaison avec les services de la CNIL et la future instance de gouvernance s'articule autour de trois axes :

a) une offre pour tous les publics en accès direct sur internet

La manière la plus simple de mettre à disposition des données anonymes issues de bases indirectement nominatives est évidemment de produire des tableaux statistiques répondant par des données agrégées aux questions que se posent les différentes catégories d'utilisateurs. Le groupe n'a pas exploré cette piste faute de temps mais elle ne pose pas de problème de principe : seulement de faisabilité technique et de coûts. La limite de l'exercice est que la demande est potentiellement infinie, alors que la production de statistiques agrégées a un coût et répond souvent mal ou trop lentement aux besoins particuliers des uns et des autres, notamment aux besoins des chercheurs ou à des questions d'une grande technicité ou exigeant le recours à plusieurs sources de données.

Une solution est de rendre plus « agile » la production de ces tableaux statistiques à l'aide d'outils techniques par exemple en mettant à la disposition des utilisateurs des bibliothèques de requêtes en libre service ; à la limite cela revient à laisser interroger la base à l'aide d'outils logiciels paramétrables fournissant des réponses anonymes sans que l'auteur de la requête ait lui-même accès aux données. Le même résultat peut être obtenu, plus classiquement en organisant des collaborations (un « guichet ») entre une équipe d'experts attachés à la base de données et les experts « métier » des demandeurs, pour fabriquer les sorties agrégées adaptées aux besoins, sachant que la mobilisation d'agents pour des travaux à façon a un coût et que là aussi on ne répond pas toujours aux besoins les plus pointus.

L'autre solution explorée par le Groupe de travail « risques de ré-identification » consiste en la production d'un grand nombre de « jeux » de données rendus anonymes, tels que pour chacun de ces jeux, le nombre de personnes présentant des caractéristiques identiques soit suffisant ($K=10$ au moins dans le test) et que parmi ces personnes il y en ait toujours un nombre suffisant ($L=3$ au moins dans le test) qui aient des maladies différentes. On a démontré que c'est possible et que les critères d'anonymat peuvent être vérifiés. Des précautions doivent être prises cependant pour qu'il ne soit pas possible de reconstituer ainsi des jeux de données identifiants. Une sécurité supplémentaire peut être obtenue en utilisant des techniques d'échantillonnage (si la règle de tirage est secrète).

Les fichiers qui seront mis à disposition de cette façon doivent être absolument anonymisés, pour tout type de consultant qui n'a d'ailleurs pas besoin de se faire connaître (assureur, employeur, voisins, famille...).

Le groupe de travail est parti du fichier PMSI et a construit, par généralisation ou suppression de variables, des jeux répondant aux critères d'anonymisation : pour toute combinaison de variables quasi identifiantes, il y a au moins 10 individus. Parmi ces individus, il y a au moins 3 types de « maladies » différentes.

À titre d'exemple, on a construit un fichier comprenant pour chaque séjour d'hospitalisation : le sexe et l'âge (en 18 groupes d'âge) de la personne hospitalisée, son mode d'entrée et de sortie

¹³ Le groupe rappelle qu'il n'est pas besoin d'être médecin pour respecter et faire respecter le secret professionnel sur des données personnelles de santé.

(en 2 groupes : « domicile » ou « autre ») et la durée d'hospitalisation (12 modalités) et le groupe homogène de malades (indicateur de la maladie).

Un autre fichier donnant un peu moins de précision sur la durée d'hospitalisation (2 modalités), mais comprenant la région de résidence du malade semble aussi répondre aux critères. D'autres exemples ont été étudiés et un grand nombre de jeux répondant aux critères d'anonymisation peut a priori être mis à disposition de la même manière,

Des fichiers analogues pourront, selon la même méthode, être construits à partir du SNIIRAM.

Le groupe ne donne pas de recommandation définie sur le nombre minimal de personnes présentant les mêmes caractéristiques que doit comporter un jeu de données pour être considéré comme anonyme (k-anonymat). Certains membres du groupes préconisaient d'examiner les seuils $K=3$ et $L=3$ qui préservent en théorie l'anonymat en admettant qu'il serait peut-être nécessaire de se fixer sur un seuil un peu plus élevé ($K=5$, $L=3$ par exemple). Les travaux devront être poursuivis à ce sujet, car certains membres du groupe considéraient à l'inverse que ces valeurs sont trop basses pour garantir l'anonymat.

Il est mis à l'étude aussi la proposition que des fichiers plus détaillés sur les caractéristiques médicales du séjour (médicaments en sus, DMI, diagnostics et actes...) soient fournis mais sans indication du nom de l'établissement et sous la forme de sondage (au 1/10 voire 1/3). Certes, les données médicales ne sont pas considérées comme un quasi identifiant (c'est l'information que « l'assaillant » est supposé *rechercher* dans la base : s'il en disposait déjà, il n'aurait pas besoin de la base) mais elles doivent être appauvries elles aussi dans les jeux de données anonymes parce que la partie médicale du fichier de résumé des sorties anonymisées (RSA) est très *discriminante* (chaque RSA est à cet égard différent des autres et on pourrait, si on conservait cette information commune dans des jeux de données par ailleurs appauvris reconstituer le jeu d'origine).

La technique de sondage confère en tout état de cause une sécurité supplémentaire et peut être appliquée partout où un fichier exhaustif n'est pas indispensable, à la condition bien sûr qu'on ne parte pas d'un fichier où chaque personne a déjà des caractéristiques uniques et que la règle de tirage soit gardée secrète..

Le caractère anonyme de ces fichiers, notamment l'impossibilité de remonter au fichier d'origine, devra être validé par des instances à désigner (on pense notamment à un avis de la CNIL sur le mode de construction de ces fichiers et non sur chaque fichier pris isolément).

La piste du « bruitage » (ajout ou substitution de données rendant « fausses » les données individuelles mais conservant des valeurs telles que moyenne, médiane, écart type etc.) n'a pas été suivie, notamment parce qu'elle nous a paru supposer une utilisation prédéterminée du jeu de données alors qu'en open data, les usages sont à la main des utilisateurs. Elle suppose également une compétence forte des utilisateurs pour en maîtriser les limites d'utilisation.

b) une offre plus limitée, destinée à un public ciblé : un nouvel EGB ?

S'agissant des données présentant un risque de présomption d'identification, la proposition faite est de constituer au moins un échantillon de données de grande taille tel qu'on ne puisse y reconnaître personne de manière certaine et dont l'accès soit ouvert à des organismes publics ou privés dont les missions et les compétences le justifient. Le principe de l'accès leur étant acquis, une limitation s'impose cependant en raison d'un risque résiduel tenant au caractère très particulier de certains parcours de soins¹⁴ : il conviendra qu'ils souscrivent un certain nombre

¹⁴ Dans l'échantillon, si la règle de tirage est secrète et si certaines données sont rendues suffisamment imprécises, on ne pourra jamais affirmer qu'une personne est reconnaissable même elle présente (âge, sexe, parcours de soins) des caractéristiques uniques.

d'engagements sur le respect des règles d'accès, d'usage des données et de publication des résultats.

Il appartiendrait à la CNIL après avis d'une instance technique de valider le niveau de risque. Dès lors, l'accès à tel ou tel organisme (chercheurs, organisations professionnelles, administrations) pourrait relever comme aujourd'hui d'un arrêté ministériel.

Un fichier spécifique pourrait être construit à partir du SNIIRAM, chaîné avec le PMSI. Ce fichier ne serait pas exhaustif, mais issu d'un sondage (par exemple au 1/10¹⁵). Il est important que le mode de sondage et les paramètres retenus pour celui-ci demeurent strictement confidentiels. Il sera alors impossible à quiconque de savoir si telle ou telle personne est présente dans ce fichier.

Si l'on découvre alors une personne ayant exactement les caractéristiques d'un individu (et d'un seul) présent dans ce fichier, on ne pourra pas en conclure à son identification, car :

- on ne sait pas si la personne visée est ou non présente dans le tirage ;
- il n'y a qu'une certaine probabilité (dans l'exemple, une chance sur dix) pour que la personne repérée dans le fichier soit la personne connue.

Cette remarque doit cependant être tempérée par le fait qu'une accumulation de ressemblances avec une personne connue pourrait, au fil de l'addition des millésimes diffusés, au fur et à mesure de l'enrichissement annuel de la base, conduire à une probabilité d'identification proche de l'unité.

C'est pourquoi, ce type de fichier devrait être :

- rendu imprécis dans certaines de ses dimensions (telles que dates et lieu de soins, adresse, âge, date de décès ?) ce qui le rendrait inutilisable pour le suivi précis des parcours de soins mais très utile pour des études de pharmaco-épidémiologie par exemple...
- mis en accès limité, avec identification de la personne qui y a accès et engagement de sa part à n'utiliser le fichier qu'à des fins déclarées à l'avance, lesquelles n'incluent pas la tentative d'identifier une personne (modèle du réseau Quetelet ou de l'accès actuel des chercheurs INSERM aux données de l'EGB).

c) une offre restreinte pour les fichiers permettant une identification indirecte des malades

Comme cela a été vu, certains fichiers, utiles pour la recherche, sont extrêmement détaillés et permettent une identification indirecte de nombreuses personnes.

Les fichiers PMSI et EGB ayant été assez largement diffusés, ils ne devraient dorénavant être disponibles que moyennant des procédures d'accès très strictes.

Les personnes demandant un accès doivent justifier de la raison qui leur fait effectuer cette demande. Elles doivent démontrer les garanties de sérieux pour elles-mêmes et pour leur environnement professionnel. Au moment de leur accès, une trace des opérations effectuées doit être conservée. Enfin, un organisme doit vérifier que les sorties effectuées à partir de ces fichiers ne mettent pas en danger la protection de la vie privée des personnes.

Des systèmes et des procédures existent qui permettent ce type d'accès à des données confidentielles sans que cela implique un site unique.

Mais si ces conditions sont réunies et qu'on admet que le fichier est indirectement¹⁶ nominatif, il n'y a sans doute pas lieu d'appauvrir les données fournies : la distinction actuelle entre les

¹⁵ Un échantillon au 1/10 a été suggéré dès 2006 dans un rapport des Professeurs Bégaud et Costagliola.

¹⁶ Bien sûr, puisque les traitements autorisés visent exclusivement à tirer des conclusions générales à partir de données individuelles, il n'y a pas lieu de mentionner dans les données traitées les données directement identifiantes (nom, NIR, adresse etc.)

personnes autorisées ou non à croiser les quatre données sensibles (dates de soins, mois de naissance, commune de résidence et date de décès) n'a a priori plus de raison d'être.

Ce système à trois niveaux devrait permettre de présenter une offre aussi vaste que possible, tout en préservant la confidentialité des données.

Annexe 10

Contributions des membres de la commission

Contribution à la commission open data en santé transmise par Chantal Cases (INED), Marcel Goldberg (INSERM/UVSQ), Yann Bourgueil (IRDES), Jean Paul Moatti (INSERM/AVIESAN), Grégoire Rey (INSERM/AVIESAN), Pierre-Yves Geoffard (CNRS-PSE Ecole d'économie de Paris)

La nécessaire séparation de la gouvernance et de l'expertise

Le projet de rapport de la Commission Open data propose, dans la partie consacrée au dispositif de régulation des accès au SNIIRAM-PMSI concernant les accès ponctuels à des données porteuses de risque de réidentification liés à des projets, un modèle d'organisation des accès comprenant deux structures séparées : une structure de gouvernance chargée d'élaborer la doctrine et de veiller à sa déclinaison opérationnelle, et une instance technique chargée de l'évaluation des demandes en liens avec des projets spécifiques.

La première, dont le rôle serait de fixer les orientations et les conditions de mise en œuvre des accès aux données de santé doit être composée de l'ensemble des parties prenantes qui ont une légitimité à faire valoir leurs intérêts : professionnels et établissements de santé, organismes gestionnaires des bases de données, chercheurs, industriels, représentants des patients, organismes d'assurance maladie complémentaire, État.

La seconde serait une instance scientifique et technique, composé d'experts des différents domaines disciplinaires concernés, chargée d'instruire les demandes d'accès en la plaçant sur le seul champ de l'expertise au regard d'un référentiel d'analyse élaboré dans le cadre d'orientations établies par l'instance de gouvernance.

En tant que chercheurs concernés par l'utilisation du SNIIRAM-PMSI ou d'autres bases de données individuelles en santé, nous soutenons fortement cette proposition de séparation formelle entre gouvernance et expertise, qui correspond aux pratiques usuelles en matière d'évaluation scientifique adoptées de façon universelle. A notre sens, ces pratiques sont également adaptées aux demandes relevant de ces accès ponctuels qui relèveraient d'autres opérateurs que des chercheurs. Nous soutenons également la mise en place de procédures simplifiées quand il s'agit de méthodologies de références, pour des données à faibles risques de ré-identification.

Cette séparation est d'autant plus justifiée lorsqu'il s'agit de travailler sur des données de santé. En effet, les données de santé à caractère personnel sont des données sensibles, et à ce titre leur traitement est par principe interdit. Seule une finalité d'intérêt collectif justifie les dérogations prévues par la loi 78-17 du 6 janvier 1978 modifiée fixant les règles applicables en matière de traitement des données individuelles. Il est donc indispensable de s'assurer que les projets s'appuyant sur l'accès à ce type de données offrent *a priori* des garanties suffisantes de bénéfice collectif, tant au niveau des objectifs de connaissance que de pertinence des traitements au regard de la finalité poursuivie, des compétences scientifiques du demandeur et de sa capacité à respecter un référentiel de sécurité et de préservation de la confidentialité des accès que de l'engagement à rendre publics les résultats obtenus. Notamment, la qualité méthodologique du projet doit être bonne, car une investigation mal faite peut conduire à des décisions erronées qui peuvent avoir un impact profond sur la santé des personnes. La qualité des projets a donc une dimension éthique qui doit être prise en compte, et un projet qui ne correspondrait pas à ces critères ne saurait justifier l'accès à des données de santé à caractère personnel, et deviendrait contraire aux règles de l'éthique scientifique et du contrat implicite passé entre la communauté

scientifique, les sujets sur lesquels porte l'investigation et le reste de la société : cette dernière accepte que des investigations de plus en plus poussées soient menées sur ses membres à la condition qu'un mécanisme de contrôle existe permettant de veiller à ce que les résultats des études puissent bénéficier à l'ensemble de la collectivité.

Nous recommandons également d'envisager que les projets de recherche ayant fait l'objet d'une évaluation positive par des comités scientifiques indépendants, notamment dans le cadre d'appels à projets, ou s'inscrivant clairement dans le programme de recherches de structures labellisées par les Alliances publiques de Recherche et ayant fait l'objet d'une évaluation indépendante par le Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur (qui succède à l'AERES) puissent faire l'objet d'une procédure simplifiée qui tienne compte de ces évaluations antérieures. Dans tous ces cas, le rôle de l'instance d'expertise scientifique et technique ne doit pas être de doubler l'évaluation proprement scientifique de la qualité du projet, mais plutôt de vérifier que les données dont l'accès est demandé correspondent bien aux objectifs proposés et que les conditions de sécurité et les procédures fixées seront bien respectées.

Dans ce contexte, il est absolument indispensable que la phase d'évaluation scientifique soit tout à fait indépendante de la gouvernance générale des accès aux données individuelles du SNIIRAM-PMSI. En effet s'il est justifié que l'instance de gouvernance soit composée des parties prenantes, car celles-ci sont légitimes à faire valoir leurs intérêts, ceux-ci ne peuvent entrer en ligne de compte dans l'évaluation scientifique et technique.

Contribution de l'ITMO santé publique



[Contribution ITMO santé publique.pdf](#)

Contribution de l'UNPS



[Contribution UNPS.pdf](#)

Contribution du Syntec-numérique



[Syntec Numérique – Position paper.pdf](#)



[Syntec Numérique – Position paper – Annexes.pdf](#)

Contribution de la Fédération Hospitalière de France

La fédération Hospitalière de France (FHF) partage les objectifs annoncés d'ouverture des données de santé.

Elle considère que c'est un facteur majeur d'amélioration de la qualité et d'efficacité économique. Cependant de nombreux freins subsistent en particulier concernant les bases de données de l'assurance maladie et dans une moindre mesure celle de l'Atih.

Afin de garantir un accès large dans un cadre sécurisé, la FHF propose que l'ensemble des données de l'assurance maladie et de l'agence technique de l'information hospitalière soit géré par une institution indépendante qui aurait en charge l'ensemble des nomenclatures, classification et échelle de coûts pour la partie scientifique.

Cet organisme regrouperait les bases de données du SNIIR-AM et du PMSI.

Cette organisation serait garante d'une véritable ouverture de données dans un cadre sécurisé.

La mise en œuvre pourrait être réalisée à moyen constant par redéploiement de moyens techniques du SNIIR – AM et de l'Atih.

Contribution conjointe du LEEM et du SNITEM



[Contribution conjointe du LEM et du SNITEM.pdf](#)

Contribution du SNITEM



[Contribution du SNITEM.pdf](#)

Contribution de la FEHAP



[Contribution FEHAP-Open Data en sante](#)

Contribution de la FHP



[Contribution de la FHP.pdf](#)

Contribution de l'IDS

Propositions d'amendements au pré rapport examiné par la commission open data en santé le 30 avril 2014

Après 6 réunions de la commission open data en santé, et après avoir écouté et lu les préoccupations exprimées par les différentes parties prenantes, un compromis est facilement accessible si ces dernières le souhaitent.

C'est la raison pour laquelle les propositions ci-dessous visent à avoir un caractère consensuel en permettant de bâtir une gouvernance des données de santé qui :

- est simple, transparente, cohérente, efficace et soucieuse des finances publiques
- permet l'expression des aspirations et des besoins en matière de données de santé de l'ensemble des parties prenantes.

1) l'importance de la place de la société civile au sein de la gouvernance des données de santé

L'ouverture des données de santé est un enjeu de société. La présence active de la société civile est donc essentielle non seulement pour animer le débat de société mais aussi pour constituer un des piliers d'une réelle démocratie sanitaire. La société civile doit en outre avoir accès aux données de santé car elle a démontré sa capacité à utiliser ce capital immatériel que constituent les bases de données pour améliorer le fonctionnement du système de santé, sans que cela porte atteinte à la vie privée des personnes. La gouvernance des données de santé – c'est-à-dire la supervision du dispositif d'ouverture des données et les autorisations d'accès - doit permettre l'atteinte de ces deux objectifs complémentaires.

2) la supervision du dispositif d'ouverture des données et les autorisations d'accès

Les amendements portent sur la deuxième option du pré-rapport examiné le 30/4 par la commission open data en santé.

En effet, comme cela a été dit par différents acteurs, la supervision du dispositif d'ouverture des données de santé et les autorisations d'accès, pour l'open data appelé « régulé » puisque concernant des données considérées comme pouvant être indirectement ré-identifiantes, doivent être rassemblées au sein d'une même entité juridique, l'IDS. C'est le scénario le plus simple, transparent, cohérent, efficace et soucieux des finances publiques. C'est aussi celui qui répond le mieux aux

aspirations de la société civile décrites ci-dessus, car il lui permet d'agir à la fois sur l'expression des besoins et les accès accordés. Enfin, sur un plan institutionnel, s'il y avait deux organismes distincts et indépendants, l'un chargé de la supervision du dispositif d'ouverture des données, et l'autre des autorisations d'accès à ces données, cela comporterait des risques évidents de conflits, et donc de blocages.

Pour autant, cette deuxième option décrite dans le pré-rapport examiné le 30/4 par la commission open data en santé doit être amendée pour répondre aux aspirations exprimées le 30/4 par toutes les parties prenantes.

2.1 Les premiers aménagements portent sur les critères d'évaluation des autorisations d'accès qu'il est proposé d'amender de la manière suivante (page 53 du pré-rapport examiné le 30/4):

- Le bénéfice attendu en matière de santé publique, de protection de la santé, de protection sociale et de gestion des services de santé, de médecine préventive, de diagnostics et de soins, et de connaissance scientifique (y compris dans les domaines économique et social) des traitements envisagés ;
- La pertinence des traitements au regard de la finalité poursuivie et le besoin effectif d'accéder aux données ;
- Les **références** données par le demandeur quant à ses compétences et à sa capacité à respecter un référentiel de sécurité et de préservation de la confidentialité des accès ;
- La transparence du demandeur sur les méthodes employées et les résultats obtenus.

Ces modifications sont apportées notamment afin d'être en phase avec l'évolution de la réglementation européenne, telle qu'elle résulte du projet de règlement européen, et éviter des contentieux futurs

2.2 Les autres aménagements portent sur les critères sur lesquels le comité technique rend son avis et la portée de son avis.

2.2.1 Les critères sur lesquels le comité technique rendrait son avis seraient :

- La pertinence des traitements au regard de la finalité poursuivie et le besoin effectif d'accéder aux données ;
- Les **références** données par le demandeur quant à ses compétences et sa capacité à respecter un référentiel de sécurité et de préservation de la confidentialité des accès ;

La composition du comité technique doit être conforme aux critères qu'il examine.

Les instances délibérantes de l'IDS (bureau et assemblée générale) examineraient les deux autres critères, leur composition étant adaptée à cet examen, puisqu'elles rassemblent les acteurs de la santé, notamment la société civile :

- Le bénéfice attendu **en matière de santé publique, de protection de la santé, de protection sociale et de gestion des services de santé, de médecine préventive, de diagnostics et de soins**, et de connaissance scientifique (y compris dans les domaines économique et social) des traitements envisagés ;
- La transparence du demandeur sur les méthodes employées et les résultats obtenus.

2.2.2 La portée de l'avis rendu par le comité technique

L'avis du comité technique, sur les critères qu'il serait chargé d'examiner, serait rendu de manière totalement indépendante (y compris par le fait que chacun de ses membres établirait une déclaration de conflits d'intérêts) et transmis aux instances délibérantes de l'IDS. Ces dernières examineraient ensuite les critères qui relèvent de leur champ de compétence. Elles approuveraient ou non les demandes d'accès avant l'autorisation de la CNIL. Elles ne pourraient s'opposer à l'avis du comité technique. Elles pourraient soit faire leurs observations, soit les signaler dans le cadre de l'approbation d'accès qu'elles donneraient ou non au demandeur.

Les motifs d'approbation ou de refus d'accès seraient rendus publics.

Différents paragraphes du pré-rapport examiné le 30/4 par la commission open data en santé doivent donc être amendés pour tenir compte de cette proposition.

3) Sur d'autres points, le pré rapport examiné le 30/4 par la commission open data en santé doit être amendé comme suit :

Page 5 Introduction

(...)

Les membres de la commission ont partagé des attentes communes et fortes. Ce sont ces attentes que ce document retrace ; les préconisations, si elles ne recueillent pas toujours un accord unanime, visent à répondre à ces attentes ; c'est à l'aune de leurs résultats que leur mise en œuvre devra être évaluée.

Ce paragraphe doit être mis au point après les débats de la dernière réunion de la commission open data en santé

Page 23

(...)

Plusieurs membres soulignent néanmoins l'avancée constituée par la création de l'Institut des données de santé (créé par la loi en 2004, et mis en place effectivement en 2007).....

Page 26

→ une base de données individuelles des bénéficiaires (DCIR) pour réaliser des études plus approfondies sur la consommation des soins : 6 types d'organismes ont accès au DCIR-PMSI (CNAMTS, CCMSA et RSI ; CNSA ; INVS ; HAS ; ANSM ; ARS sur leur champ de compétence régional et sous l'autorité d'un médecin salarié de l'agence).

Page 27

→ accès temporaires : l'IDS approuve les accès temporaires (données agrégées, échantillon EGB) pour les organismes publics à des fins d'études en santé publique : **54** accès temporaires sont ouverts à **30** organismes.

3.5.2 L'expression des parties prenantes (page 52)

(...)

Les membres de la Commission insistent sur le fait **qu'une instance permettant l'expression des acteurs ne trouve** sa justification qu'à la condition qu'elle soit dotée d'une véritable capacité à **pouvoir agir à la fois** sur les orientations et **sur** la mise en œuvre des accès aux données de santé. **Pour ce qui concerne l'expression des besoins des acteurs**, il pourrait être proposé que les ministres chargés de la sécurité sociale et de la santé transmettent chaque année au Parlement un rapport sur la politique de mise en œuvre du dispositif, et que ce rapport soit précédé d'un avis de cette instance. **Pour ce qui concerne les autorisations d'accès, afin de pouvoir agir réellement sur elles, l'instance serait en charge de les approuver après avis indépendant d'un comité technique (cf. selon les modalités décrites ci-dessus)**

3.5.3.2 Les accès liés à des projets (page 54)

(...)

Par ailleurs, la Commission est favorable à la demande émise de longue date **par les acteurs de la santé d'utiliser toutes les possibilités offertes par la loi du 6 janvier 1978 pour l'emploi du NIR afin de faciliter les appariements de données, notamment des données du SNIIRAM avec des données individuelles recueillies auprès des patients dans le cadre d'enquêtes ou de suivi de cohortes.**

Les dispositions combinées du IV de l'article 26 et du III de l'article 27 de la loi du 6 janvier 1978 permettent de procéder par un « *acte réglementaire unique* » à l'autorisation de tous les traitements qui « *répondent à une même finalité, portent sur des catégories de données identiques et ont les mêmes destinataires ou catégories de destinataires* ». Cette possibilité a été mise en œuvre pour la première fois dans le cadre d'un décret du 9 novembre 2012. A cette occasion, la section sociale du Conseil d'Etat a estimé que « *la finalité en vue de laquelle des traitements de données peuvent être autorisés par un acte réglementaire unique doit être définie de manière*

suffisamment précise par cet acte et correspondre à un objectif suffisamment homogène et circonscrit ». L'acte réglementaire unique ne peut donc porter sur une finalité définie de manière trop générale.

3.5.4 Un niveau opérationnel de mise à disposition des données (page 55)

La mise en place d'une organisation facilitant les accès est essentielle sauf à créer un droit virtuel. **L'ATIH et la CNAMTS responsables** d'organiser ces accès, doivent satisfaire à un cahier des charges garant de la qualité du service rendu au demandeur et du respect de la sécurité et de la confidentialité des données. Un service d'extraction de données devrait être également offert **au sein de l'ATIH et de la CNAMTS** incluant la mise à disposition de données rendues anonymes. Ce service à façon pourrait s'inscrire dans un modèle économique fondé sur la facturation des prestations concernées.

Ces prestations doivent être servies par un service dédié et identifiable **au sein de l'ATIH et de la CNAMTS** afin notamment de permettre à l'instance de gouvernance (décrite en 3.5.2) d'évaluer la qualité du service rendu et d'en publier les résultats.

Bien entendu, la centralisation et la consolidation des données issues des systèmes d'information de gestion continueront d'être réalisées par les organismes (ATIH et CNAMTS) qui en sont chargés aujourd'hui. Les représentants des professionnels de santé et des établissements de santé doivent continuer d'être associés au suivi de ces procédures dans le cadre d'un comité de gestion.



Contribution de l'UNOCAM

1. Open data en santé : un enjeu pour l'efficacité de notre système de santé

L'UNOCAM soutient pleinement la démarche initiée en faveur de l'ouverture des données publiques de santé, y compris celles concernant les établissements et professionnels de santé, en raison des enjeux majeurs qui y sont associés (et exposés dans la partie 2 du rapport) :

- Renforcer la démocratie sanitaire
- Renforcer l'autonomisation du patient
- Accroître l'efficacité de l'action publique
- Favoriser la recherche, l'innovation et le développement industriel.

Pour les organismes complémentaires d'assurance maladie, l'open data contribuera à améliorer l'efficacité globale de notre système de santé :

- l'accès aux données permet de construire un partage des connaissances entre acteurs du système de santé (AMO, AMC et offreurs de soins) et représentants de ces usagers fondé sur une base d'informations commune qui servira donc de référence et facilitera le dialogue entre les parties prenantes. Ce partage des connaissances est indispensable pour le débat démocratique, l'évaluation de la réponse aux besoins de santé et la recherche collective d'une plus grande efficacité du notre système de santé.
- l'accès aux données, de nature et d'origine diverses, permettra à chaque acteur du système, quel que soit son rôle (offreur de soins, régulateur, financeur...), d'améliorer l'efficacité de son action, par une meilleure connaissance et un plus grand partage des informations disponibles.
 - o C'est le cas pour l'action publique pour laquelle la transparence pourrait contribuer à un décloisonnement entre institutions (caisses d'assurance maladie, Etat, agences sanitaires, agences régionales de santé, acteurs locaux...) favorable à l'efficacité collective.
 - o C'est le cas pour les professionnels de santé et acteurs de terrain, en renforçant la logique du retour d'information (qui, pour mémoire, constituait l'un des objectifs fondateurs du SNIIRAM) et en l'élargissant à une meilleure connaissance globale du système de soins.

- C'est le cas pour les organismes complémentaires d'assurance maladie : en tant que financeurs et porteurs du risque pour la part de dépense sur laquelle ils interviennent, l'accès permanent à une information statistique plus large et plus riche que les seules données de leur propre portefeuille est essentiel. Cette information leur permet de mieux apprécier les risques qu'ils couvrent, d'accroître les possibilités de régulation et de réduction des restes à charge, d'améliorer les services rendus aux assurés.
- enfin, la pluralité des utilisations de ces données enrichira la documentation de ces bases et poussera à l'amélioration de leur qualité et de leur capacité à répondre aux besoins des acteurs du système de santé et à ses usagers.

Pour les organismes complémentaires d'assurance maladie, l'amélioration de l'information du citoyen sur les professionnels et établissements de santé, les parcours de soins, les médicaments et les traitements, etc., est indispensable. Si le préalable consiste à travailler sur les besoins des citoyens, la pertinence et l'actualisation des données rendues publiques, la pédagogie nécessaire à leur diffusion, il ne doit pas constituer un frein à l'ouverture de ces données et leur réutilisation. L'accès effectif à l'information du citoyen doit être garanti par la publicité faite sur cet accès mais aussi par les possibilités de démultiplication de l'information et de réutilisation par des tiers.

2. Accès aux données : les acteurs à pied d'égalité

Pour l'UNOCAM, si l'ouverture des données publiques (agrégées ou individuelles) à tous les acteurs, publics comme privés, est souhaitable, cette ouverture doit se faire dans le cadre de la protection de la vie privée. L'UNOCAM est favorable au renforcement de l'effectivité des sanctions, en cas de manquement à ce principe fondamental.

Dès lors que les données sont parfaitement anonymes, leur ouverture et leur réutilisation doivent être organisées à court terme. L'UNOCAM soutient les propositions faites en ce sens en ateliers.

Pour les données présentant un risque potentiel de réidentification, l'UNOCAM soutient 3 types de mesures complémentaires :

- premièrement, la poursuite des travaux relatifs à l'évaluation du risque, de façon à mesurer si le risque est réel et, le cas échéant, à caractériser son niveau ;
- deuxièmement, une action sur ces données, destinée à supprimer le risque de réidentification et permettre un premier niveau d'ouverture : production de tableaux statistiques, exclusion des données sensibles, micro-agrégation, floutage, renforcement des critères d'anonymat par le relèvement du niveau minimum de personnes présentant des caractéristiques identiques... ;

- troisièmement, la régulation de l'accès aux données exhaustives (le risque de réidentification ne devant pas entraîner la fermeture des accès alors que les besoins d'utilisation de données fines et complètes, par différents acteurs du système, sont réels). Le pré-rapport stipule qu'hormis les services de l'Etat (pour lesquels la décision relève de la compétence du Ministre de la santé), la décision d'autoriser tel ou tel organisme, de façon permanente ou de manière ponctuelle sur un projet, est fondée sur l'analyse de ses missions/de son projet, et sur les garanties apportées en termes de respect de la confidentialité, d'organisation des accès et le respect d'un référentiel de sécurité. Si l'UNOCAM adhère à ce principe général, elle souhaite qu'il soit assorti de règles claires et non discriminantes :
 - o Les organismes publics ou privés doivent être traités à égalité dès lors qu'il s'agit de demandes pouvant intéresser la santé publique ou la gestion du risque santé. Les conditions d'accès doivent donc, a priori, être identiques, sans discrimination selon le statut du demandeur.
 - o Pour l'examen des missions/projets, si un critère de bénéfice collectif devait être retenu, il ne pourrait être réduit à la seule amélioration de la connaissance scientifique ou de santé publique, mais être étendu à l'amélioration de la connaissance économique et sociale. La pluralité des critères concourant au bénéfice collectif doit être reconnue.
 - o S'agissant des garanties à apporter, un cahier des charges devra être établi à cet effet.

3. Une gouvernance simple et garante des principes d'ouverture des données de santé

Pour l'UNOCAM, la gouvernance de l'open data en santé doit reposer sur 2 principes :

- o La simplification des démarches et la lisibilité des process, comme élément de promotion de l'open data ;
- o La pluralité des points de vue, comme garant du respect des principes posés pour l'accès aux données de santé (cf. ci-dessus) ;

C'est pourquoi l'UNOCAM défend l'idée de renforcer le rôle de l'Institut des données de santé (IDS) dans la gouvernance de l'accès aux données de santé :

- La composition de l'IDS serait adaptée de façon à ce que l'ensemble des parties prenantes y soient représentées : Etat, opérateurs publics, caisses d'assurance maladie, organismes d'assurance maladie complémentaires, professionnels et établissements de santé, chercheurs, industriels, représentants des patients et des usagers, etc. ;

- L'IDS serait l'instance chargée de l'instruction et de l'autorisation des accès permanents et ponctuels dans le cadre de l'open data « régulé ». Dans un souci d'efficacité, l'IDS serait structuré autour de deux instances : un comité technique (chargé de l'expertise des dossiers, sur la base des principes et critères définis, transmise à la CNIL) et un conseil de surveillance (lieu d'échange et de partage de la connaissance entre acteurs et garant du respect des procédures d'octroi des accès aux données de santé). Cette organisation offre l'avantage d'avoir une structure unique et un interlocuteur unique.

En conclusion, l'UNOCAM soutient pleinement la démarche initiée en faveur de l'ouverture des données publiques de santé. Les travaux de la Commission posent de bons principes basés sur une réelle volonté de partage et de transparence, dans un cadre sécurisé mais aussi simplifié. Comme de nombreux acteurs, l'UNOCAM appelle à la concrétisation très prochaine de ces grands principes et au lancement des travaux qui en découlent. Elle considère que cette démarche contribuera à améliorer l'efficacité globale de notre système de santé.

En conséquence, l'UNOCAM restera attentive à **deux questions majeures** associées à la mise en œuvre effective des intentions inscrites dans le projet de rapport de la commission : la gouvernance et le modèle économique retenu.

Sur **la question de la gouvernance**, l'UNOCAM considère qu'il est indispensable d'y associer de façon effective les composantes de la société civile dans toute leur diversité, dans une instance indépendante. Sans nier le rôle central de l'Etat et de la CNIL en tant que garants de la protection de la vie privée, la gouvernance doit être réellement partagée afin de reconnaître la pluralité des enjeux associés à l'ouverture des données de santé et d'évaluer, collectivement, les bénéfices et les risques associés à la mise à disposition des données, en définissant, le cas échéant, les règles à respecter pour les accès à des données avec risque de réidentification. C'est pourquoi l'UNOCAM défend l'idée de renforcer le rôle de l'Institut des données de santé (IDS) dans la gouvernance des accès.

Sur **la question du modèle économique**, l'UNOCAM estime qu'il est urgent de traiter des moyens et sources de financement alloués à l'ouverture des données de santé. De ce modèle dépend en effet l'effectivité de la mise en œuvre des principes posés par la commission, tant en termes d'accès immédiat à l'information et de réutilisation de celle-ci, que d'efficacité dans le processus de mise à disposition des données aux acteurs intéressés (chercheurs, acteurs du système de santé, professionnels, usagers...).

CONTRIBUTION DE CHRISTIAN BAUDELLOT ET CHRISTIAN SAOUT

Objet : Réflexions suite aux échanges lors de la réunion du 2 avril 2014.

Le pré rapport proposé aux membres de la commission open data en santé le 2 avril 2014 traduit des évolutions pour une plus grande ouverture des données de santé, quel que soit le statut juridique des acteurs. Le point de vue de la société civile, s'il n'est pas encore suffisamment pris en compte, semble avoir été pour partie retracé.

C'est ainsi qu'ont été bien distingués :

- l'Open Data en santé stricto sensu, c'est-à-dire les données qui devraient être accessibles directement, y compris par le grand public, car complètement anonymisées,
- l'Open Data que l'on peut appeler « régulé », qui devrait concerner uniquement des données anonymisées considérées comme pouvant éventuellement présenter un risque de ré-identification des personnes. Leur mise à disposition devrait faire l'objet d'une instruction individuelle. Il ne s'agit pas d'interdire leur utilisation, quel que soit le statut juridique du demandeur, mais de les mettre à disposition sous certaines conditions (voir plus bas).

Par ailleurs, tout au long des travaux de la commission, a aussi été mise en évidence l'existence de nouvelles technologies permettant d'éviter des risques éventuels de ré-identification des personnes en raison d'un éventuel acte malveillant répréhensible juridiquement. La micro-agrégation permet par exemple, à partir de données anonymisées considérées comme potentiellement ré-identifiantes, d'agréger des variables (mois et année de naissance en lieu et place du jour, mois et année de naissance par exemple) pour les rendre ainsi complètement anonymisées sans pour autant les appauvrir afin de répondre aux besoins. Ce travail doit pouvoir être engagé sans tarder afin de pouvoir mettre à disposition de tous les acteurs, y compris le grand public, dans le cadre de l'Open Data stricto sensu, un grand nombre de données anonymisées.

*

Des échanges qu'ont pu avoir postérieurement à la réunion du 2 avril dernier de nombreux membres, représentant la société civile, il ressort qu'il est nécessaire d'opérer encore des avancées.

A cet égard, le rapport de la commission devrait indiquer aussi clairement, en ce qui concerne l'Open Data « régulé » :

- Que les organismes publics et privés seront traités à égalité dès lors qu'il s'agit de demandes pouvant intéresser la santé publique ou la gestion du risque maladie.

- Que les demandes d'accès directs permanents ou ponctuels, ou d'extractions de données, relevant de l'Open Data « régulé », feront l'objet d'une instruction individuelle sur la base d'un cahier des charges (garanties apportées en termes de respect de la confidentialité, régulation des personnes habilitées, traçabilité des accès et des requêtes etc.).
- Qu'une instance de régulation élaborera le cahier des charges que la CNIL approuve.
- Qu'une instance de régulation approuvera les accès directs permanents et ponctuels, ou les extractions de données, relevant de l'Open Data « régulé », sauf dans deux types de cas :
 - les organismes en ayant une manifeste et impérieuse nécessité pour l'exercice de leurs missions (les services de l'Etat et ses opérateurs directs). La décision relèvera de la compétence directe du Ministre en charge de la santé,
 - les organismes de toute nature ayant besoin pour répondre à des demandes faites et validées par une autorité publique (HAS, ANSM, ...), ces demandes n'ayant pas lieu d'être réinstruites.
- Que l'instance de régulation comprendra un organe technique sur lequel elle s'appuiera en tant que de besoin pour élaborer le cahier des charges et instruire les demandes.
- Que tous les accès directs permanents et ponctuels, ou les extractions de données, relevant de l'open data « régulé », feront l'objet d'une autorisation de la CNIL, après l'approbation de l'instance de régulation (ou du Ministre en charge de la santé dans les cas qui relèvent de lui).

*

En ce qui concerne la gouvernance du dispositif, de nombreuses voix considèrent que pour mettre en œuvre de manière opérationnelle l'Open Data « régulé » la société civile doit y être pleinement associée, dans une gouvernance simple et centrée autour de l'Institut des données de santé. A cet égard, les avancées, qu'il faut reconnaître au terme des échanges du 2 avril dernier, resteront un vœu pieu si la gouvernance des accès reste dans un entre soi administratif ou dans une instance nouvelle dénuée de pouvoir et de compétences.

Ainsi, dès lors que les membres de la commission Open Data ont souligné tout au long des travaux la pertinence de l'action de l'Institut des données de santé depuis 2007 pour plus de connaissance, de partage, de transparence, de déontologie à travers la charte qu'il a élaborée, et d'équité entre les acteurs de la santé, c'est bien autour de cet institut qu'il faut consolider ce qui a déjà fait ses preuves et qui marche. L'Institut des données de santé ayant acquis à ce titre, expérience et légitimité.

Dès lors, conformément à la loi de 2004, l'Institut des données de santé doit être l'instance de régulation qui à la fois :

- instruit et autorise les demandes d'accès permanents et ponctuels, ou les extractions de données, relevant de l'Open Data « régulé »,
- et supervise le dispositif décrit ci-dessus et ses évolutions.

Ces deux missions ne sauraient être disjointes pour les confier à deux organes différents comme l'envisage le pré-rapport remis le 2 avril. Une telle gouvernance bicéphale est inefficace, source de confusions et opaque. A rebours, elles doivent être liées pour plus d'efficacité et de simplicité, et éviter des conflits futurs. Le pré rapport du 2 avril 2014 évite une source de conflits en repositionnant, conformément à la loi, les missions du COPIIR SNIIRAM. Il faut éviter de recréer une autre source de conflits potentiels. Un organe unique est aussi la solution la plus économique et la plus claire pour tous les acteurs.

Dans un tel schéma « unifié », la gouvernance et les moyens de l'Institut des données de santé peuvent bien entendu être revus afin notamment de distinguer en son sein :

- d'une part, une instance composée, comme aujourd'hui, des principaux acteurs, qui instruit et autorise les accès permanents et ponctuels dans le cadre de l'Open Data « régulé », avec en tant que de besoin l'avis d'un comité technique
- d'autre part, une sorte de Conseil de Surveillance, à créer, qui rassemblerait un panel le plus large possible d'acteurs et qui superviserait le dispositif d'accès aux données de santé.

En revanche, la mise à disposition des données relevant de l'Open Data « régulé » doit rester de la compétence des gestionnaires de ces bases, notamment de la CNAMTS et de l'ATIH qui ont toute la légitimité et les compétences requises.

Enfin, la mission du COPIIR SNIIRAM doit être recentrée sur la gestion et le renseignement de ce système d'information conformément à la loi de financement de la sécurité sociale pour 1999.

Contribution de l'UNPS

OPEN DATA ET DONNEES MEDICALES INDIVIDUELLES

Dans le cadre de l'Open Data, dans une volonté de faire vivre la démocratie sanitaire, se pose la problématique d'une publication de données de santé liées à l'activité personnelle des praticiens. Si sur le principe il est difficile de s'opposer à l'amélioration des pratiques médicales et donc à la publication de ces données, il existe toutefois un danger important dans ces publications tant au niveau initial par celui qui corrige les données médicales, que par l'utilisation qui peut en être faite par des tiers sans possibilité pour le praticien d'avoir accès ni aux sources, ni à l'utilisation ultérieure, avec un risque réel de publication ou d'utilisation erronée.

Il faut souligner que déjà l'interprétation des données du PMSI par la presse, conduit à l'élaboration de palmarès dont les critères sont parfois loin d'être pertinents, conduisant des établissements à être positionnés d'une année à l'autre d'une manière radicalement différente en bien ou en mal sans possibilité de réel regard sous la pertinence et l'application des critères utilisés pour juger la qualité médicale de l'établissement. Une telle application au niveau individuel du médecin aurait des effets délétères inacceptables.

Il faut opposer la méthodologie de certification des établissements de la Haute Autorité de Santé, à celle de l'Assurance Maladie pour les médecins. Dans le cadre de la certification des établissements, les cadres administratifs et médicaux de l'établissement sont confrontés pendant cinq jours aux experts-visiteurs, puis un pré-rapport est élaboré par la H.A.S auquel l'établissement peut répondre point par point évitant ainsi toutes mauvaises interprétations, omissions ou malentendus. C'est une démarche remarquable qui permet alors une publication objective des données de qualité de l'établissement tant localement que sur les sites de l'HAS. Cette méthodologie a incontestablement conduit à améliorer l'organisation administrative et médicale des établissements.

En ce qui concerne l'Assurance Maladie, il n'en n'est pas de même puisque l'agrégation et l'interprétation des données se fait exclusivement au sein des services de l'Assurance Maladie, régulièrement lors des négociations conventionnelles nous sommes amenés à redresser des présentations qui nous sont faites et qui ne correspondent pas d'une manière exacte à la réalité de l'organisation au quotidien. Ceci pour de multiples raisons et en particulier la difficulté d'identifier par exemple la qualification du médecin fournissant les données. A titre d'exemple dans le domaine vasculaire, les actes peuvent être faits par des médecins spécialistes cardiologues, par des médecins angiologues, parfois par des médecins spécialistes en médecine générale, mais également des chirurgiens vasculaires ou des radiologues. Il est alors difficile pour l'Assurance Maladie de préciser avec pertinence qui est responsable de la pratique de tel ou tel acte. Les conséquences sont minimes puisqu'il s'agit de données collectives et qu'il n'y a aucune publication vis-à-vis de l'extérieur ou simplement parfois des présentations générales dans le cadre de conférences de presse.

Une étape complémentaire a été faite avec la mise en ligne des honoraires sur le site Ameli-direct. Cette mise en ligne se fait le plus souvent à l'insu du praticien, le principe n'est pas attaquant, par contre sa pertinence est parfois très loin de la réalité. D'une part seul un certain nombre d'actes

figurent sur le site de l'Assurance Maladie, d'autre part, notamment en cas de changement d'organisation tarifaire, les erreurs sont nombreuses sans compter les bugs informatiques ou les tarifs ont été multipliés par cent y compris en terme de remboursement, pour une difficulté de maîtrise informatique du monde Apple. Nous sommes très loin du recueil et de l'utilisation conjointe établissement HAS.

L'autre élément à prendre en compte est l'utilisation par des tiers des données de l'Assurance Maladie déjà compliquée lorsqu'il s'agit de publier les tarifs. Cette situation pourrait devenir totalement inacceptable si les données de rémunération sur objectif de santé publique était prise en compte dans la volonté de l'Open Data. En effet la ROSP est basée sur un certain nombre de critères élaborés dans le cadre conventionnel mais qui doit prendre en compte la situation du praticien et de sa patientèle. Ces critères sont forcément différents en fonction de la patientèle avec des pertinences variées selon que le praticien exerce dans le 16^{ème} arrondissement de Paris ou en Seine-Saint-Denis. Des comparaisons hâtives entraîneraient forcément un préjudice individuel pour le praticien et ne contribueraient pas à la démocratie sanitaire. Nous sommes très loin dans le cadre de la ROSP, de la démarche de la H.A.S et surtout le praticien éprouve de grandes difficultés à faire corriger les données qui lui sont attribuées liées à la lourdeur administrative et technologique du recueil de l'Assurance Maladie.

Le dernier écueil à la publication de données de type ROSP est leur utilisation ultérieure par des tiers. Face aux difficultés des praticiens vis-à-vis de l'Assurance Maladie pour obtenir un correctif des données qui lui sont attribuées, leur utilisation par des tiers rendrait impossible toute mise à jour puisque l'Assurance Maladie elle-même ne serait plus maître de ceux qui utilisent les données qu'elle publie d'une manière erronée.

C'est donc avec réticence que nous voyons certaines volontés mises en place dans ce rapport, si les intentions de transparence ne sont pas contestables, le mode de réalisation reste particulièrement difficile l'exemple récent du dossier médical personnel, de son échec, nous laisse extrêmement interrogatifs sur la réalisation d'une publication des données médicales des praticiens.