

LA PLACE DE LA SCIENCE DES DONNEES
AU SEIN DE L'IDS Institut Droit et Santé (Faculté de Droit, Université Paris
Descartes)

Présentation de l'IDS

L'Institut droit et santé (IDS) a été fondé en 2005 il est co-dirigé par Mme Anne LAUDE (Doyen de la Faculté de Droit) et M. Didier TABUTEAU. Il s'agit d'un laboratoire labellisé INSERM qui comprend une vingtaine d'enseignants-chercheurs et le même nombre de doctorants.

L'Institut droit et santé réunit une équipe pluridisciplinaire regroupant des juristes (privatistes et publicistes), des médecins, des pharmaciens et un sociologue. Ce caractère pluridisciplinaire se trouve encore renforcé par les partenariats noués avec l'extérieur (par exemple avec la Chaire santé de Science-po).

Il nous a été demandé, au cours de cette journée, d'intervenir pour présenter :

- 1) le besoin scientifique et la dimension big data de nos problèmes (étendue ou diversité de la base de données, rapidité de calcul et algorithmes, persistance, respect de la vie privée, sécurité, ...)
- 2) les possibilités d'applications interdisciplinaires connues
- 3) les besoins en infrastructure (stockage, cpu)
- 4) les liens avec la formation
- 5) les laboratoires et/ou autres structures concernées
- 6) la participation à des grands programmes nationaux
- 7) les liens industriels, européens, internationaux

Cette intervention sera surtout consacrée au premier point, soit à la dimension big data de nos recherches.

L'IDS s'intéresse bien évidemment à la science des données et ce à un double titre : par la constitution et l'utilisation de bases de données pour son activité de recherche (I), par la réflexion sur l'utilisation des données de santé (II).

I.- Constitution et utilisation de bases de données pour l'activité de recherche de l'IDS

L'IDS mène de nombreuses études sur les évolutions du droit de la santé, ce qui suppose évidemment l'utilisation de certaines bases de données.

Etude relative à la judiciarisation de la santé (La judiciarisation de la santé, par A Laude, J. Parriente et D.Tabuteau, Éditions de Santé, 2012).

Une des études récentes très importante que l'on peut citer est celle sur la judiciarisation de la santé, publiée en 2012.

La recherche a été réalisée à partir de l'analyse des décisions de justice recueillies auprès des juridictions administratives, des cours d'appel et de la Cour de cassation, de données globales transmises par le Ministère de la justice et relatives à l'activité des TGI et des juridictions disciplinaires. L'analyse de l'activité des juridictions en matière de responsabilité des professionnels de santé a été couplée à l'analyse des données de l'ONIAM, des assureurs des professionnels, des établissements de santé, de l'AP-HP et du Médiateur de la République.

Les bases de données constituées par l'IDS à cet effet ont été croisées entre elles et à d'autres sources pour enrichir l'analyse (INSEE, DREES notamment).

Cette étude a été menée à partir de 50 000 décisions de justice rendue de 1999 à 2009.

La méthodologie employée afin de construire les bases de données s'est essentiellement appuyée sur le recueil des données concernées.

A cette fin, des membres de l'IDS se sont **en premier lieu** déplacés au Centre de recherches et de diffusion juridiques du Conseil d'Etat. Un travail de recueil de décisions de justice a été réalisé à partir de la base Ariane, qui regroupe toutes les décisions rendues par les tribunaux administratifs (TA), cours administratives d'appel (CAA) et le Conseil d'Etat (CE) pour la période 1999-2009.

Cette méthode a permis de recueillir et d'examiner 21 483 décisions administratives rendues entre 1999 et 2009. Toutes n'ont pas été considérées comme pertinentes, 7000 ont été finalement retenues

En deuxième lieu, des membres de l'IDS se sont rendus au Service de documentation et d'études de la Cour de cassation. Un travail de recueil des décisions de justice a été réalisé à partir des bases Jurica et Jurinet. La base Jurica regroupe les décisions des cours d'appel (CA) pour la période 2006-2009.

Une méthode similaire à celle employée pour effectuer les recherches au Conseil d'Etat a été utilisée. Elle a permis de recueillir 26 789 décisions de cours d'appel rendues entre 2006 et 2009.

Après exclusion des décisions hors champ de la responsabilité des professionnels de santé et hors période d'étude, ce sont près de 1 300 décisions qui ont été retenues pour figurer dans les bases IDS relatives au contentieux des cours d'appel et de la Cour de cassation.

Pour l'AP-HP, l'ONIAM et les assureurs, les données relatives aux indemnisations définitives sur la période 2006-2009 sont disponibles auprès de l'Observatoire des risques médicaux (ORM).

Cette étude a permis d'identifier, de quantifier et d'analyser les données conduisant à apprécier la réalité et la portée de la judiciarisation de la santé sur la décennie écoulée et de relativiser la portée du phénomène.

Autres études

D'autres études qui utilisent des bases de données et qui ont donc un rapport avec la science des données sont aussi actuellement en cours, comme une en particulier sur une comparaison du circuit des plaintes France-Québec. L'objectif de l'étude est de dresser une évaluation des plaintes au sein des établissements de santé en France et au Québec pour apprécier si les cadres juridiques sont adaptés. Des bases de données relatives aux circuits des plaintes seront ainsi constituées pour mener cette étude.

II.- Réflexion sur l'utilisation des données de santé

D'autre part et c'est sur ce point qu'il faut insister plus longuement, l'IDS s'intéresse beaucoup à la question de l'encadrement juridique des données de santé et, plus particulièrement du Big Data et de l'Open Data Santé.

En effet, compte tenu du développement et de l'importance croissante des nouvelles technologies en matière de santé, notre équipe s'est fixée pour objectif d'engager des recherches dans **l'axe de recherche « Technologies et santé »**. Un des centres d'intérêt spécifique de cet axe concerne les multiples questions, de plus en plus cruciales, relatives au **traitement et à la sécurité de l'information en santé**.

Au travers notamment des exemples de la télémédecine et des réseaux en santé, nos recherches essaient plus généralement d'analyser l'articulation des règles du code de la santé publique et du code de la sécurité sociale avec la loi informatique et libertés de 1978 quant à la **sécurité des « données de santé à caractère personnel »**, reconnues comme données sensibles en 2004.

Les questions de l'Open data et du Big Data Santé nous intéressent particulièrement.

Sur le premier point, l'Open Data Santé, l'IDS suit avec attention l'examen de l'article 47 du projet de loi Santé (Article 193 du texte adopté n° 650, 17 décembre 2015, Projet de loi de modernisation de notre système de santé)

L'article 47 est présenté par le gouvernement :

« Cet article pose les bases nécessaires à la mise en œuvre d'un dispositif d'accès aux données de santé.

Le Gouvernement s'est prononcé en faveur de l'ouverture des données en santé et a chargé une commission de lui faire des propositions en ce sens. Ce travail est à la base de cet article, qui entend réformer l'accès aux données de santé afin que leurs potentialités soient utilisées au mieux dans l'intérêt de la collectivité, et du principe de valeur constitutionnelle de protection de la santé, tout en assurant la

confidentialité des données personnelles, qui procède du droit au respect de la vie privée, autre exigence de rang constitutionnel, d'autant plus forte qu'il s'agit de données sensibles.

La conciliation à assurer entre ces principes est aujourd'hui rendue complexe par la diversité et l'importance des enjeux sanitaires, démocratiques et économiques ainsi que par une gouvernance éclatée des traitements de données de santé et par des règles manquant de clarté.

Afin de répondre à ce constat, les dispositions de l'article créent dans le code de la santé publique un système national des données de santé (SNDS) qui centralisera les données des bases existantes en matière sanitaire et médico-sociale et assurera leur mise à disposition selon deux modalités distinctes, présentant chacune des règles d'accès garantissant un niveau de protection des données adapté à leur nature. D'une part, les données pour lesquelles aucune identification n'est possible seront accessibles et réutilisables par tous, en open data. D'autre part, les données potentiellement identifiantes (bien qu'elles ne comportent ni les noms et prénoms ni le numéro de sécurité sociale), pourront être utilisées sur autorisation de la CNIL à des fins de recherche, d'étude ou d'évaluation d'intérêt public dans le domaine de la santé, ou sur autorisation par décret en Conseil d'État après avis de la CNIL pour l'accomplissement des missions de service public, à des conditions rigoureuses assurant la protection de ces données sensibles.

La gouvernance du système s'appuiera sur un Institut national des données de santé réunissant les parties prenantes dans un groupement d'intérêt public. Dans le cadre d'orientations fixées par l'État, la gestion opérationnelle de la base de données sera confiée pour l'essentiel à la Caisse nationale de l'assurance maladie des travailleurs salariés.

Cet article prévoit aussi un régime de déclaration à la CNIL et non plus d'autorisation pour les traitements de données personnelles à réaliser en urgence en cas d'alerte sanitaire.

Il modifie les dispositions de l'article L. 1111-8-1 du code de la santé publique pour permettre l'utilisation du numéro d'inscription au répertoire des personnes physiques (NIR ou numéro de sécurité sociale), qui a l'avantage d'être un identifiant opérationnel et certifié, déjà connu des professionnels, comme identifiant des personnes prises en charge dans le champ sanitaire et médico-social, l'utilisation restant cantonnée à ces champs. Il servira dans les faits essentiellement à la bonne tenue des dossiers médicaux.

L'article réforme et clarifie enfin le régime d'autorisation des traitements de données personnelles, qu'elles soient ou non issues du SNDS, à des fins de recherche, d'étude et d'évaluation en matière de santé, en fusionnant et adaptant les chapitres IX et X de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés. Il soumet ces traitements, y compris s'ils nécessitent l'usage du NIR pour apparier des données de sources différentes, à l'autorisation de la CNIL, qui recevra sur le plan technique un éclairage d'un comité d'expertise (ou d'un comité de protection des personnes pour les recherches intéressant la personne humaine). L'Institut national des données de santé aura également pour rôle d'éclairer la décision de la CNIL sur l'intérêt public des recherches, études et évaluations ».

Un Colloque sur le Projet de loi santé a ainsi été organisé le Mercredi 12 novembre 2014, à la Faculté de droit. **M. Pierre-Louis Bras** (Inspecteur général des affaires sociales, professeur associé à l'Université Paris Descartes, membre de la Chaire santé de Sciences Po) auteur du rapport sur la gouvernance et les données de santé en date de 2013 est intervenu à ce sujet. **Mme Lydia Morlet**, MCF et membre de l'IDS, a présenté plus particulièrement les modifications législatives concernant le DMP (Travaux à paraître).

Par ailleurs, l'IDS a organisé deux séminaires sur les applis santé, qui nous ont donné l'occasion de réfléchir sur les bases de données susceptibles d'être constituées à partir des données produites par les smartphones.

- Séminaire Applis Smartphones et santé, Promesses et menaces, sous la direction de Mme A. Laude, 26 juin 2014, publié dans le Journal de Droit de la Santé et de l'Assurance Maladie (JDSAM), 1^{er} décembre 2014.
- Séminaire Applis Smartphones et Santé, sous la direction de Mme A. Laude et de Monsieur G. Duhamel, 9 et 10 juin 2015, Centre Universitaire des Saints-Pères, Travaux à paraître

Le Big Data santé concerne aussi les données produites par les objets connectés (c'est-à-dire les objets dont la connexion à internet n'est pas la finalité principale tels les bracelets, montres, scanners de poche ou des piluliers intelligents...).

Une fois les données -produites soit par les smartphones, soit par les objets connectés - collectées, en effet, se pose la question de **l'exploitation des bases de données** ainsi constituées. Une interrogation concerne aussi la loi susceptible d'être appliquée au traitement de ces données. Apple avec son HealthKit, Samsung avec SAMI, Google avec Google Fit prévoient de stocker sur une plateforme une série d'informations comme le poids ou le rythme cardiaque, la tension..., transmis par le téléphone en lien avec différents objets connectés. Si ces données concernent des utilisateurs européens, la question du respect du droit de l'Union peut se poser.

Ensuite, plusieurs finalités peuvent être assignées à ces bases de données : une finalité immédiate, celle pour laquelle elles ont été constituées : s'assurer qu'une personne se nourrit, est bien à son domicile, qu'elle prend bien ses médicaments... Au regard de cette première finalité, se posent essentiellement des questions liées à la sécurisation de ces bases contenant tant de données relatives à la vie privée des personnes.

Une deuxième finalité susceptible d'être envisagée concerne la recherche. En effet, de nombreux enseignements sont susceptibles d'être tirés de ce type de base. Il s'agit de bases de données comportementales qui pourraient éventuellement intéresser les chercheurs, même si la segmentation de ces bases, par opérateur proposant des objets connectés rend, à l'évidence, celles-ci incomplètes. En outre, l'exploitation secondaire de ces bases pour déterminer des habitudes de consommation par des opérateurs privés devrait être strictement encadrée.

La question de la conformité de cette exploitation secondaire avec la loi informatique et libertés doit être posée. En outre, l'ensemble des questions relatives à la protection des données doit être examiné au regard du **futur règlement européen** en cours d'adoption.