Intelligence artificielle et données de santé

Lorraine Maisnier-Boché

Intelligence artificielle et santé

Lorraine Maisnier-Boché
Juriste en droit des nouvelles technologies, Asip Santé (l'agence française de la santé numérique)

Intelligence artificielle et données de santé

L'intelligence artificielle est profondément liée aux données, en ce qu'elles constituent à la fois la condition pour que l'intelligence artificielle se développe et l'objet sur lequel l'intelligence artificielle intervient.

L'amélioration de l'intelligence artificielle passe en effet par l'éducation des machines, et celle-ci s'opère principalement par l'exploitation de bases de données. L'intelligence artificielle se nourrit des techniques de *big data*, tout particulièrement en analysant massivement des données permettant d'identifier des personnes physiques (internautes, clients, prospects, utilisateurs…), c'est-à-dire des données à caractère personnel.

Le régime juridique qui encadre l'utilisation de ces données constitue ainsi un élément à prendre en compte pour exploiter une intelligence artificielle, *a fortiori* lorsque celle-ci utilise des données à caractère personnel de santé, données sensibles au sens de la loi n° 78-17 du 6 janvier 1978, relative à l'informatique, aux fichiers et aux libertés, modifiée, dite loi « Informatique et Libertés ».

Si ces modalités d'utilisation des bases de données sont nouvelles, et rejoignent par exemple la problématique du *big data*, elles sont d'ores et déjà appréhendées par le droit de la protection des données à caractère personnel de santé **(I)**.

Par ailleurs, la façon dont une intelligence artificielle raisonne et prend une décision est également encadrée par le droit positif, bien que des réflexions soient en cours pour en améliorer l'effectivité **(II)**.

I. Eduquer l'intelligence artificielle : les conditions d'utilisation des bases de données de santé

Dans le domaine de la santé, le cadre juridique applicable à de telles bases a été façonné par les principes issus de la loi Informatique et Libertés et du code de la santé publique.

Si ces règles protectrices ont vocation à s'appliquer aux bases contenant des données de santé à caractère personnel **(A)**, elles peuvent également s'appliquer à des bases de données en principe « anonymisées » **(B)**.

A - Cadre juridique de l'utilisation des bases de données de santé à caractère personnel

L'entrainement d'une intelligence artificielle grâce à une base de données à caractère personnel constitue bien évidemment un traitement de données couvert par les dispositions de la loi Informatique et Libertés.

Malgré la nature disruptive des technologies d'intelligence artificielle et les progrès significatifs que celles-ci ont effectués depuis quelques années[1], les principes fondamentaux de la loi Informatique et Libertés qui, dès 1978, plaçait l'informatique au service de chaque citoyen, demeurent pleinement d'actualité[2].

Les garanties apportées aux personnes concernées sont adaptées aux traitements réalisés par l'intelligence artificielle, que ce soit le droit d'être informé préalablement de l'existence d'un tel traitement, le droit d'accéder aux données traitées voire le droit de s'y opposer pour un motif légitime[3]. Ces garanties sont aujourd'hui renforcées par la consécration d'un principe d'autodétermination informationnelle, selon lequel les personnes ont le droit de décider et de contrôler l'usage qui est fait de leurs données à caractère personnel[4].

A cet égard, il faut cependant relever que l'intelligence artificielle transforme la place de la personne dans ce nouveau marché que constituent désormais les données à caractère personnel. Il était d'usage de considérer qu'un service gratuit impliquait que l'utilisateur de ce service était le produit. A présent, l'utilisateur ne constituerait qu'une donnée permettant d'éduquer une intelligence artificielle (« *training data* »). Traditionnellement, les données à caractère personnel d'un individu possédaient une valeur économique propre, en raison de la possibilité pour un responsable de traitement de les vendre pour démarcher les personnes, pour établir des profils, etc. A présent, leur utilisation aux fins de nourrir une intelligence artificielle ne leur reconnaitrait au niveau individuel qu'une valeur marginale, dès lors qu'elles sont traitées de façon massive. La capacité pour les personnes de reprendre le contrôle sur leurs données pourrait en être amoindrie.

1 - Rapport d'information n° 4594, de C. de Ganay et D. Gillot, établi au nom de l'office parlementaire d'évaluation des choix scientifiques et technologiques, « *Pour une intelligence artificielle maitrisée, utile et démystifiée* » p.31

2 - E. Geffray, « *Quelle protection des données personnelles dans l'univers de la robotique ?* », Dalloz IP/IT, juin 2016, p.295

3 - Art. 32, 38, 39 et 40 de la loi Informatique et Libertés

4 - Art. 1 de la loi Informatique et Libertés, modifié par la loi n°2016-1321 du 7 octobre 2016 pour une République Numérique. Voir notam. Y. Poullet et A. Rouvroy, « *Le droit à l'autodétermination informationnelle et la valeur du développement personnel* », Etat de droit et virtualité, Thémis, 2009 ; N. Botchorichvili et L. Maisnier-Boché, « *Loi République numérique : état des lieux en matière de protection des données personnelles ?* », RLDI, 2016/131.

Par ailleurs, le régime juridique spécifiquement applicable aux données à caractère personnel de santé implique certaines contraintes concernant l'utilisation de bases de données de santé. La sensibilité de ces données restreint les fondements juridiques permettant d'effectuer un traitement, qui sont limitativement énumérés par la loi Informatique et Libertés[5]. L'information préalable qui doit en tout état de cause être fournie aux personnes nécessite de préciser les finalités pour lesquelles les données seront traitées, ce qui comprend les finalités pour lesquelles l'intelligence artificielle doit être entrainée. Dans la négative, l'utilisation de la base à cette fin pourrait être considérée comme constituant un détournement de finalité. De surcroît, le traitement massif de données de santé par l'intelligence artificielle impose de réaliser préalablement une étude d'impact sur la vie privée des personnes concernées, conformément aux dispositions du nouveau Règlement général sur la protection des données[6].

Enfin, les dispositions du code de la santé publique relatives au secret professionnel prévoient que toute personne prise en charge par un professionnel, un établissement ou une structure du secteur sanitaire, médico-social ou social a droit au respect de sa vie privée et du secret des informations la concernant[7]. Il faut rappeler que l'obligation au secret professionnel est par principe absolue et que personne, pas même le patient, ne peut en affranchir le professionnel[8].

La diffusion des données de santé n'est pas libre, l'échange ou le partage de ces données devant se faire dans le respect du secret professionnel, en circonscrivant les cercles dans lesquels les données peuvent être communiquées, avec le consentement préalable du patient ou sous réserve de son information préalable, le patient pouvant exercer son droit d'opposition[9]. Ainsi, la mise en place d'une intelligence artificielle au sein d'un ou plusieurs établissements de santé, utilisant les dossiers médicaux comme données de référence, devrait respecter ces principes d'échange et de partage, notamment entre équipes de soins distinctes… Enfin les traitements aux fins de recherche, d'étude ou d'évaluation dans le domaine de la santé sont soumis à un complexe régime d'autorisation[10].

Malgré ces nombreuses dispositions, certaines questions persistent quant à la compatibilité de ce cadre juridique avec l'intelligence artificielle.

De façon similaire aux traitements fondés sur le *big data*, le principe de minimisation des données, qui impose de ne traiter que les données strictement nécessaires au regard de la finalité du traitement, peut sembler contradictoire avec la nécessité d'éduquer l'intelligence artificielle grâce à des volumes conséquents de données. De même, la finalité des traitements opérés par l'intelligence artificielle ne serait pas nécessairement identifiée au moment de la collecte[11]. En effet, l'intelligence artificielle apprend de ces données et en tire des opérations et réponses pour lesquelles elle n'avait pas été programmée initialement. Ceci impacte également l'information qu'il est possible de donner aux personnes dont les données sont traitées.

Il est possible de relativiser cette crainte vis-à-vis de l'intelligence artificielle : contrairement aux traitements de *big data*, dans lesquels des tendances ou corrélations sont recherchées au sein de grands ensembles de données sans que l'objet recherché soit identifiable, une intelligence artificielle demeure un programme, structuré par des algorithmes ayant des objectifs précis, bien que les réponses ne soient pas nécessairement prévues. Aussi, il semble envisageable dans de nombreux cas de circonscrire une finalité aux traitements utilisant une intelligence artificielle. Encore faut-il que les personnes soient informées de cette nouvelle finalité et le cas échéant, aient pu consentir au traitement.

B - Bases de données anonymisées et risque de réidentification

Par principe, anonymiser une base de données permet de s'affranchir du régime de protection des données à caractère personnel, dès lors qu'une fois anonymisées, les données n'identifient plus une personne physique et ne constituent donc plus des données à caractère personnel.

A titre d'illustration, le système national des données de santé, qui symbolise la démarche d'*open data* de santé, revêt ainsi deux formes : des jeux de données sous forme de statistiques agrégées, accessibles par tous, et des jeux de données sous une forme potentiellement identifiante, dont l'utilisation est soumise à l'autorisation de la CNIL après avis du Comité d'Expertise pour les Recherches, les Etudes et les Evaluations dans le domaine de la Santé (CEREES) et le cas échéant de l'Institut National des Données de Santé (INDS) et d'un comité de protection des personnes[12]. L'anonymisation de certains jeux de données permet ainsi leur mise à disposition au public, sans passer par un régime complexe d'autorisation, tout en préservant les impératifs de la protection des données à caractère personnel.

Cependant, les affaires AOL, Netflix et State of Massachussetts ont révélé qu'une base de données en principe anonymisée peut contenir des informations

5 - Art. 8 de la loi Informatique et Libertés

6 - Art. 35-2 du Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (Règlement général sur la protection des données)

7 - Art. L.1110-4 du code de la santé publique

8 - Cass. crim. 5 juin 1985, n°85-90322

9 - Art. L.1110-4 et L.1110-12 du code de la santé publique

10 - Chap. IX de la loi Informatique et Libertés et titre II du Livre 1 de la première partie du code de la santé publique

11 - A. Bensamoun et G. Loiseau, « *L'intégration de l'intelligence artificielle dans certains droits spéciaux* », Dalloz IP/IT, mai 2017, p.295

12 - Art. 193 de la loi n° 2016-41 du 26 janvier 2016 de modernisation de notre système de santé, et décret n° 2016-1871 du 26 décembre 2016 relatif au traitement de données à caractère personnel dénommé « Système National des Données de Santé »

suffisamment précises pour identifier un individu, ou bien peut être croisée avec une autre base pour réidentifier les personnes concernées[13].

Depuis lors, le niveau d'exigence retenu par le groupe de travail de l'article 29 (G29) et la CNIL, récemment confirmé par le Conseil d'État, invite à la prudence en ce qui concerne les procédés d'anonymisation, en raison de ce risque de « ré-identification » ou de « dé-anonymisation ».

Une base de données anonyme est celle qui ne permet pas ou plus d'identifier une personne physique dès lors que suffisamment d'éléments lui ont été irréversiblement retirés, et qui fait l'objet d'un suivi tel que la réidentification de la personne physique demeurera impossible[14]. Dès lors qu'une donnée correspond à une seule personne physique, la CNIL semble considèrer qu'il ne s'agit pas d'une donnée anonymisée mais simplement pseudonymisée : un processus d'anonymisation doit empêcher « *d'isoler un individu dans un ensemble de données, de relier entre eux deux enregistrements dans un ensemble de données (ou dans deux ensembles de données séparés) et de déduire des informations de cet ensemble de données* »[15]. Le niveau exigé pour qu'une base soit qualifiée d'anonyme apparait ainsi particulièrement difficile à atteindre, sauf à se cantonner à des bases de statistiques descriptives[16].

C'est la raison pour laquelle certains considèrent que l'anonymisation est fondamentalement contraire à l'utilité des données, et qu'afin de conserver à une base de données une utilité pour ceux qui veulent l'analyser, l'anonymisation doit impérativement être imparfaite : « *une donnée peut être soit utile, soit parfaitement anonymisée, mais elle ne peut jamais être les deux* »[17]. La perte de la richesse informationnelle d'une base de données impliquée par un procédé d'anonymisation est-elle compatible avec la sophistication des intelligences artificielles actuelles, qui nécessitent des données d'entrainement toujours plus riches ou plus nombreuses ?

Cet affaiblissement de la richesse informationnelle des bases anonymes ne conduira-t-elle pas à une compensation via une augmentation du volume de données collectées, et partant, à une intrusion plus importante dans la vie privée des personnes ?

13 - P. Ohm, « *Broken promises of privacy : responding to the surprising failure of anonymization* », 57 UCLA Law Review, 2010, p.1717

14 - Avis du groupe de travail de l'article 29 n°5/2014 (WP216) sur les techniques d'anonymisation

15 - CNIL, délib. n°2015-255 du 16 juillet 2015 ; CE, 8 février 2017, n° 393714 : Comm. com. électr. 2017, comm. 37, n° 4, N. Metallinos ; R. Perray, J. Uzan-Naulin, « *Existe-t-il encore des données non personnelles ?* », Dalloz IP/IT, mai 2017, p.286. Voir également CNIL, délib. n°2017-145 du 9 mai 2017)

16 - L. Maisnier-Boché, « *L'anonymisation des données à caractère personnel : que faire pour sortir de l'impasse ?* », Expertises des systèmes d'information, n°416, sept. 2016, p.296

17 - P. Ohm, *op. cit.* p.1704 : « *Data can be either useful or perfectly anonymous but never both*.»

II. Exploiter l'intelligence artificielle : transparence et contrôle des algorithmes

Le cadre juridique de l'utilisation des bases de données de santé est aujourd'hui clairement établi. L'enjeu corollaire est d'assurer une véritable transparence des algorithmes utilisés par l'intelligence artificielle, afin de conserver la capacité de contrôler cette dernière, tant pour le responsable de la machine que pour le patient concerné **(A)**. Les principes de ce cadre juridique de la transparence des algorithmes sont d'ores et déjà fixés, mais font encore face à certains défis pour garantir leur effectivité **(B)**.

A – Cadre juridique du fonctionnement de l'intelligence artificielle

Le rôle de l'intelligence artificielle a été encadré dès les origines de la loi Informatique et Libertés. Dans sa version actuelle, celle-ci interdit la prise de décision produisant des effets juridiques sur le seul fondement d'un traitement automatisé, sans intervention humaine, dès lors que ce traitement a vocation à définir le profil ou évaluer la personnalité d'un individu[18]. Ces dispositions traduisent la réticence du législateur français et européen à laisser à une machine, aussi intelligente soit-elle, le soin de prendre seule une décision sans vérification par un être humain ou sans possibilité pour la personne visée de présenter son point de vue.

Le corollaire de cette interdiction est le droit pour les personnes de connaitre les modalités de raisonnement de la machine. Traditionnellement, ce droit était uniquement prévu par la loi Informatique et Libertés. Celle-ci permet en effet aux personnes physiques d'obtenir toute information permettant de « *connaître et de contester la logique* » d'un traitement automatisé à l'origine d'une décision individuelle produisant des effets juridiques à leur égard[19]. Le Règlement général sur la protection des données reprend ce droit d'accéder aux informations relatives à l'existence d'une prise de décision automatisée, à la logique sous-jacente du traitement ainsi qu'aux conséquences prévues de celui-ci pour la personne concernée[20]. Ce droit apparait cependant restreint par le renvoi à l'article 22 du Règlement, qui cantonne la notion de décision individuelle automatisée à celle « *produisant des effets juridiques* » ou « *affectant de manière significative* » la personne.

En sus, la loi pour une République numérique octroie aux administrés ayant fait l'objet d'une décision basée sur un traitement algorithmique le droit d'obtenir communication des « *règles définissant ce traitement, ainsi que les principales caractéristiques de sa mise en œuvre* », notamment le degré et le mode de contribution du traitement algorithmique à la prise de décision, les données traitées et leurs sources,

18 - Art. 10 de la loi Informatique et libertés ; art. 22 du Règlement général sur la protection des données

19 - Art. 39, I, 5° de la loi Informatique et libertés

20 - Art. 15 du Règlement général sur la protection des données

les paramètres de traitement et, le cas échéant, leur pondération, appliqués à la situation de l'intéressé, ainsi que les opérations effectuées par le traitement[21]. Bien que ce droit spécifique apparaisse faire quelque peu double emploi avec les dispositions de la loi Informatique et Libertés, il permet ainsi aux administrés personnes morales de bénéficier de cette transparence algorithmique. En outre, la décision en cause doit comporter une mention explicite informant l'administré que cette décision est prise sur le fondement d'un traitement algorithmique, et qu'il bénéficie du droit d'accès suscité[22]. La transparence des algorithmes apparait ici comme une déclinaison de l'obligation de motivation des actes administratifs.

Le respect des principes de protection de la vie privée dès la conception (« *privacy by design* ») constitue également un moyen d'encadrer le fonctionnement des intelligences artificielles[23]. La démarche France Intelligence Artificielle recommande par exemple de promouvoir la recherche sur les modèles mathématiques et informatiques de protection de la vie privée, afin d'intégrer cet aspect dans les algorithmes permettant l'éducation et le fonctionnement des intelligences artificielles utilisant des données à caractère personnel[24]. Une garantie additionnelle pourrait être de promouvoir l' « auditabilité » dès la conception (« *auditability by design* »), à savoir de développer les algorithmes d'intelligence artificielle de façon à ce qu'ils soient compréhensibles et analysables par des tiers, afin de contrôler leur comportement et les fondements de leurs décisions[25]. Comme exemple de cette démarche, la certification des logiciels d'aide à la prescription (LAP) et d'aide à la dispensation (LAD) portée par la Haute Autorité de Santé illustre les possibilités de contrôler la conformité de tels outils à certaines exigences minimales (sécurité, efficience et bonnes pratiques)[26].

B – Le défi de l'effectivité

La prohibition posée par la loi Informatique et Libertés de la prise de décision par une machine cède notamment dès lors qu'une personne physique intervient dans le processus de prise de décision. Néanmoins, la question de l'effectivité de cette garantie se pose, compte tenu de la place croissante de l'intelligence artificielle au quotidien, notamment dans le domaine de la santé.

Il est en effet à craindre que la personne chargée d'intervenir dans le processus de prise de décision ne soit pas en position réelle de contester la recommandation de l'intelligence artificielle, soit parce qu'elle ne possède pas l'expertise ou le temps d'en vérifier les fondements, soit parce que l'organisation hiérarchique de la structure en cause la contraigne en réalité à exécuter les recommandations automatisées. Pour cette raison, le Conseil d'État avait recommandé dès 2014 de prendre des mesures pour assurer l'effectivité de l'intervention humaine[27]. Il était notamment suggéré d'identifier des critères concernant la compétence de la personne qui prend la décision, la marge de manœuvre dont elle dispose ainsi que de recommander la mise en place d'un suivi du nombre de décisions prises en fonction ou en contradiction avec la proposition de l'algorithme.

Néanmoins, quand bien même les recommandations de la machine seraient systématiquement analysées par une personne capable de les contester, par exemple un professionnel de santé ayant la responsabilité de se prononcer sur le risque vital, pourrait-on lui reprocher de ne pas avoir suivi les recommandations de l'intelligence artificielle ?

Ainsi, la prohibition posée par la loi Informatique et Libertés peut paraitre assez théorique, étant donné qu'une intervention humaine ne garantit pas que la décision n'ait pas été, en pratique, bel et bien prise par la machine, dont la position ne sera pas remise en cause[28].

Et s'il est rare en pratique que la machine prenne seule une décision, en revanche nombreux sont les algorithmes de recommandations, qui influent sur les contenus et les informations qui sont présentées, et ont ainsi un impact déterminant sur la décision finalement prise par un être humain. Or, les dispositions visant à la transparence des algorithmes risquent de ne pas garantir une vision complète du raisonnement d'une intelligence artificielle. En effet, la particularité majeure de l'intelligence artificielle, à savoir sa capacité à apprendre, réduit l'intérêt d'avoir accès aux caractéristiques de fonctionnement de l'algorithme. L'intelligence artificielle est variante, en fonction des données qui lui permettent continuellement de modifier son raisonnement et ses recommandations. La frontière traditionnelle entre l'algorithme et la base de données qu'il traite tend à s'effacer[29]. Comment permettre une pleine transparence du fonctionnement d'une intelligence artificielle, constamment modifiée par son environnement ?

Enfin, le fait que les intelligences artificielles soient éduquées et modifiées en fonction des données et des cas particuliers auxquels elles sont confrontées entraine un risque qu'elles ne reproduisent la subjectivité, les inexactitudes voire les biais issus de ces bases de données[30].

21 - Art. L. 311-3-1 et R. 311-3-1-1 du code des relations entre le public et l'administration

22 - Art. R. 311-3-1-1 et svt. du code des relations entre le public et l'administration

23 - Art. 25 du Règlement général sur la protection des données

24 - France Intelligence Artificielle, Rapport de synthèse des groupes de travail, mars 2017, p.29

25 - Information Commissioner's Office (ICO), « *Big data, artificial intelligence, machine learning and data protection* », mars 2017, n° 190 et svt.

26 - Article L. 161-38 du code de la sécurité sociale

27 - Conseil d'Etat, « *Etude annuelle 2014 - Le numérique et les droits fondamentaux* », La Documentation française, 2014, p.299

28 - A. Debet, J. Massot, N. Metallinos, Informatique et Libertés, La protection des données à caractère personnel en droit français et européen : Lextenso éditions, 2015, n°853 p.378

29 - Conseil général de l'Economie, Rapport sur les modalités de régulations des algorithmes de traitement des contenus, décembre 2016

30 - Rapport au nom de l'office parlementaire d'évaluation des choix scientifiques et technologiques, « *Pour une intelligence artificielle maitrisée, utile et démystifiée* », *ibid*. p.192

Un exemple extrême est celui des algorithmes de suggestions dans les moteurs de recherche, qui peuvent renvoyer à des résultats à caractère raciste ou choquant, reflétant les nombreuses requêtes des internautes[31]. De même, le risque que les tendances détectées par l'intelligence artificielle dans un ensemble de données fassent une confusion entre corrélation et causalité a été largement démontré[32]. Que de tels biais puissent se créer dans des intelligences artificielles utilisées dans le domaine de santé et justement censées réduire le risque d'erreur humaine, apparait problématique et doit inviter à une réflexion[33].

Lorraine Maisnier-Boché

31 - ICO, *ibid*. n° 116 et svt.

32 - ICO, *ibid*. n° 118

33 - Voir les travaux de la CNIL à ce sujet : communiqué du 23 janvier 2017, « *Les algorithmes en débat* » https://www.cnil.fr/fr/ethique-et-numerique-les-algorithmes-en-debat-0. Voir également les travaux de l'INRIA sur une plateforme d'évaluation de la responsabilité et de la transparence des algorithmes « TransAlgo » : https://www.inria.fr/actualite/actualites-inria/transalgo