



Colloque : IA et mégadonnées, comment vont-elles révolutionner la recherche et la pratique médicales de demain ?

Intervention de Marie-Laure Denis, présidente de la CNIL

Vendredi 10 février 2023, Conseil d'État

Seul le prononcé fait foi

Monsieur le Vice-Président,

Madame la Présidente du collège de la HAS,

Monsieur le Président adjoint de la Section de l'Intérieur,

Monsieur le Directeur général de l'IHU ICAN,

Monsieur le Rapporteur général de la Section Sociale du Conseil d'État,

Mesdames et Messieurs,

Permettez-moi de remercier le Conseil d'État d'avoir associé la CNIL à l'organisation de ce colloque consacré à l'intelligence artificielle, appliquée aux données de santé pour la recherche et la pratique médicale de demain. Je vous remercie également pour l'invitation qui m'est faite de pouvoir ouvrir les travaux de ce colloque au Conseil d'Etat qui, une fois de plus, témoigne de son intuition pour identifier les grands enjeux, juridiques et réglementaires.

L'intelligence artificielle se généralise. Son usage se démocratise dans tous les secteurs, dans les organismes publics comme dans les entreprises privées. Il devient quotidien, y compris pour le grand public, qu'il en soit conscient ou non, grâce l'amélioration constante de la puissance d'apprentissage des ordinateurs. L'IA se propage à grande vitesse dans notre système de santé au sein duquel les données constituent un formidable levier d'innovation ainsi qu'un outil précieux de pilotage et de gestion. Des systèmes d'IA pour l'aide à la décision y sont déjà déployés et sont opérationnels, avec le développement des outils d'aide au diagnostic en imagerie médicale, d'aide à la prescription et à la décision thérapeutique.

Tous ces outils sont conçus à partir de données de santé, et même d'énormes volumes de données de santé. Ils s'en nourrissent pour leur amélioration permanente. Comme vous le savez, les données de santé ne sont pas des données personnelles comme les autres. Parce que qu'elles touchent à ce qu'il y a de plus intime chez les individus, les données de santé sont juridiquement qualifiées de « sensibles » et bénéficient, à ce titre, d'une protection renforcée. C'est aussi en raison de cette sensibilité particulière que le législateur français a choisi de

maintenir, dans ce seul domaine du droit de la protection des données, un régime de formalités préalables (de simples déclarations ou des véritables autorisations).

Ce régime peut être vécu comme contraignant pour nos chercheurs et pour les entreprises qui veulent créer et entraîner de l'intelligence artificielle à partir de données de santé. Je le sais, je ne le nie pas. Mais il a aussi sa raison d'être, ses vertus, ses avantages, que je voudrais rappeler. Et, il présente, d'ores et déjà, des flexibilités que nous devons exploiter. Ce sera le premier point que je développerai.

Par ailleurs, ce sera mon second et principal message ce matin : à la CNIL, nous avons commencé à adapter nos cadres d'analyse et notre pratique décisionnelle aux particularités de l'IA. Nous sommes prêts à continuer et à inventer, avec les professionnels, un modèle à la fois efficace et protecteur pour l'utilisation de données personnelles par l'IA.

Tout d'abord, je voudrais rappeler que la réglementation des données personnelles encadre les traitements de données de santé et de recherche mais ne les empêche nullement.

Si les principes que la réglementation fixe sont exigeants, c'est parce qu'ils ont vocation à garantir les droits fondamentaux des personnes ainsi que la fiabilité des traitements. Ces principes reposent sur des règles simples et de bon sens, comme l'information préalable du patient ou la sécurité des données, qui sont connues du monde médical.

La gestion de la crise sanitaire a démontré la robustesse de cette réglementation, puisqu'elle a permis la création, en urgence, de nombreux outils, impliquant le traitement d'important volumes de données de santé. Dans ce contexte, la conception des nouveaux systèmes d'information sanitaires et leur calibrage ont soulevé des questions inédites en matière de protection des données. Cela résultait de l'ampleur des données collectées puisque certains systèmes d'informations ont reposé sur la collecte des données d'une majorité de la population française, ainsi que sur la possibilité d'y accéder, ouverte à une multitude de professionnels. Je pense, par exemple, au SI centralisant les cas-contacts.

Cette situation exceptionnelle, caractérisée par une crise sanitaire sans précédent, n'a donc pas fait obstacle à l'application de la réglementation de la protection des données, ni à la mise en œuvre de nouvelles technologies pour des solutions innovantes comme l'application TousAntiCovid et le passe sanitaire. Le déploiement de ces outils a fait l'objet environ d'une cinquantaine de contrôles de la CNIL et, il nous est arrivé de demander les corrections nécessaires en cas de mauvaises pratiques. La prise en compte de la protection des données personnelles a été un facteur, déterminant je crois, pour contribuer à la confiance des citoyens, créer les conditions de l'acceptabilité sociale des solutions proposés et ainsi garantir, in fine, l'efficacité des outils déployés.

Au-delà du précédent que constitue cette crise, la réglementation des données de santé doit aussi être perçue comme un facteur de qualité et de fiabilité des données qui, en définitive, contribue à la valeur des résultats obtenus et participe de l'excellence de la recherche en France.

La CNIL est bien armée pour mettre en œuvre ces règles. Son expérience lui permet de cerner les cas d'usage, d'en maîtriser les problématiques, de disposer de ressources dédiées pour accompagner au mieux la conformité des acteurs du secteur de la santé. À cet égard, je veux souligner que le système des formalités préalables, permet à la CNIL d'aider les acteurs à sécuriser les traitements de données sensibles. Je pense notamment aux méthodologies de référence (MR) qui sont utilisées par la recherche ou encore au référentiel applicable aux

entrepôts des données de santé (EDS) utilisés à des fins de recherches, d'études ou d'évaluations dans le domaine de la santé. Je précise que ces outils, qui doivent être élaborés en concertation avec le secteur, peuvent, lorsque cela s'avère nécessaire, évoluer en fonction des besoins et des pratiques.

Par ailleurs, dans le secteur de la santé, la prise en compte de la dimension éthique du numérique doit favoriser un usage responsable des technologies et garantir que l'innovation se déploie au service de l'individu. La CNIL ne manque jamais de le rappeler, que ce soit dans ses avis au Gouvernement sur les projets de textes contenant des dispositions sur les données personnelles (mon espace santé, DMP, SNDS, SI COVID) ou lorsqu'elle est auditionnée par les commissions parlementaires.

Certes, la CNIL délivre aussi ce message au niveau européen. Elle joue ainsi un rôle significatif dans l'élaboration des lignes directrices adoptées par les CNIL de l'UE réunies au sein du CEPD, notamment les lignes directrices recherche. Nous avons aussi contribué activement à l'avis sur le projet de règlement sur l'espace européen des données de santé (EHDS). En définitive, la réglementation est exigeante mais elle est aussi, parfois, caricaturée. Elle n'empêche nullement l'utilisation de l'IA dans le domaine de la santé. Notre rôle, c'est donc de créer un cadre réglementaire et des méthodes d'accompagnement adaptés au développement de l'IA. Ce sera mon second point.

La CNIL accompagne les traitements à base d'IA depuis plusieurs années et nous pouvons et nous allons continuer à adapter nos instruments juridiques pour créer un cadre adapté aux acteurs de l'IA en santé, qui soit efficace et protecteur.

Cette dernière décennie, la CNIL a accompagné la mise en œuvre du système national des données de santé (SNDS), créé en 2016. Ce système unique en Europe, voire dans le monde, agrège les principales bases de données de santé en France. Depuis 2017, année de la délivrance de la première autorisation à un entrepôt de données de santé, la CNIL accompagne la massification des données de santé.

Le premier rapport éthique de la CNIL, intitulé « Comment permettre à l'Homme de garder la main ? », a permis de poser deux principes fondateurs pour une IA au service de l'homme. Le principe de loyauté selon lequel l'intérêt des utilisateurs des algorithmes doit primer, non seulement du point de vue individuel, en tant que consommateur, mais également du point de vue collectif, en tant que citoyen. Et le principe de vigilance qui préconise l'organisation d'un questionnement régulier, méthodique et délibératif sur les traitements algorithmiques. À partir de là, notre philosophie est d'accompagner l'émergence de projets conformes à ces principes éthiques. En novembre dernier, la CNIL a mis en ligne un webinaire sur l'IA et les données personnelles. Elle vient d'organiser, le 7 février, un webinaire sur les demandes d'autorisations en santé. Cette année, nous allons également publier un guide de la recherche, des recommandations sur la constitution de bases de données pour l'IA, des fiches concernant l'IA en santé et ajouter à notre cours en ligne (MOOC) un module sur la santé.

J'ai aussi voulu modifier plus radicalement les méthodes d'accompagnement de la CNIL en créant un « bac à sable ». Il s'agit d'un dispositif où on sélectionne sur un thème des projets que nous accompagnons, pas à pas, pour garantir leur conformité RGPD, pour mettre en œuvre le *privacy by design* à un stade précoce du développement du projet. Et pour inaugurer ce bac à sable en 2021, j'ai choisi le sujet des données de santé. Nous avons notamment accompagné le projet du CHU de Lille et de l'équipe Magnet de l'Inria qui concerne l'apprentissage fédéré

en intelligence artificielle appliquée aux études cliniques, ou encore, le projet Magellan, du bureau d'études Clinityx, qui facilite l'accès et l'utilisation des données du SNDS.

Je veux aller encore plus loin : trop souvent des projets sont freinés par des prétendus obstacles liés au RGPD. Ce sont parfois des obstacles fantasmés, et quand les difficultés sont réelles il y a, souvent, des solutions. Dans une dizaine de jours, nous allons donc lancer un programme d'accompagnement renforcé, une sorte de bac à sable permanent et non thématique, qui permettra à des entreprises sélectionnées de bénéficier de l'aide de la CNIL pour se développer ou monter en quelques mois un projet innovant tout en assurant sa conformité au RGPD. On peut imaginer que certains de ces projets reposeront sur l'IA dans le domaine de la santé.

L'accompagnement est une chose mais j'ai conscience que les règles elles-mêmes doivent prendre en compte le développement de l'IA, pour nous prémunir contre de nouveaux risques sans pour autant freiner l'innovation technologique. Cette position est celle de l'Union européenne qui met en œuvre une stratégie en matière de données visant à favoriser la circulation et le partage des données, tout en leur garantissant un haut niveau de protection. Dans cette perspective, la proposition de règlement IA (RIA) vise à promouvoir une IA digne de confiance, centrée sur l'humain, sûre et inclusive. Ce projet de texte propose également d'améliorer la gouvernance, de prendre en compte les droits fondamentaux et d'appliquer aux systèmes d'IA des exigences de sécurité. Ce nouveau dispositif devrait concerner les applications ayant un rôle dans la prise de décision en matière de la santé qui sont considérées à haut risque par le projet de règlement.

Ce volontarisme rejoint les propositions que le Conseil d'Etat a formulées en 2022 au Premier ministre dans son étude intitulée « Intelligence artificielle et action publique : construire la confiance, servir la performance ». Le Conseil d'Etat a notamment plaidé pour la mise en œuvre d'une stratégie de déploiement volontariste d'une IA de confiance et, compte tenu de la nécessaire articulation de l'IA avec le RGPD, a proposé de donner un rôle central à la CNIL en la matière. Je peux d'ores-et-déjà vous assurer que la CNIL partage pleinement cette analyse et les propositions qu'elle porte. Le rapporteur de cette étude, le Président Thierry Tuot, va d'ailleurs la présenter dans un instant.

Il nous faut donc adapter nos cadres d'analyses et nos outils à l'IA. Nous avons commencé : nous avons adopté, le 7 octobre 2021, un référentiel pour les entrepôts de données de santé par tous les établissements publics, par exemple les IHU. Autrement dit, nous avons défini une méthode standard pour constituer de tels entrepôts qui sont l'oxygène de la recherche en IA. Et, si cette méthode est respectée, on peut créer l'entrepôt par simple déclaration d'un « engagement de conformité » à la CNIL, sans procédure d'autorisation. C'est une première, et c'est un premier pas, récent : il faut que les acteurs s'en saisissent. S'il faut l'adapter, nous l'adapterons ; à condition que le cadre qu'il définit soit suffisamment protecteur pour les personnes, car cela reste notre boussole.

Ces pistes pourront, entre autres sujets, être débattues et enrichies à l'occasion de la table ronde de la fin de la matinée, à laquelle participera la membre du Collège de la CNIL, en charge de la santé et qui sera modérée par un ancien SG de la CNIL. Nos méthodologies de référence, notamment, sont des outils souples, qui pourront peut-être, si c'est nécessaire, accueillir de nouveaux cas d'usage correspondant à l'intelligence artificielle. Il nous faut aussi créer de la confiance et on pourrait réfléchir à développer des certifications spécifiques, garantissant le respect du RGPD, pour les dispositifs médicaux utilisant l'IA. Ce ne sont que des pistes, des réflexions que nous avons, et qui pourront nourrir vos échanges durant cette journée.

J'ajoute enfin que j'ai très récemment annoncé la création d'un service de l'intelligence artificielle au sein de la CNIL. Il s'agit de faciliter la compréhension du fonctionnement des systèmes d'IA (SIA), de consolider l'expertise de la CNIL dans la connaissance et la prévention des risques pour la vie privée liés à la mise en œuvre de ces systèmes et de préparer l'entrée en application du RIA.

Vous l'aurez compris, je suis profondément attachée à l'idée que protéger les données personnelles, ça n'est pas handicaper notre recherche médicale et notre capacité d'innovation. J'ai bien conscience que les technologies qui émergent, avec l'intelligence artificielle, sont infiniment plus consommatrices de données que celles du passé, que ce qui prévalait lorsque les principes de la loi Informatique et libertés ont été forgés. La CNIL n'est pas dans un réflexe de peur ou de blocage. À l'inverse, il n'est pas question pour nous de renoncer à la belle mission qui est la nôtre : contribuer, en appliquant des principes juridiques qui demeurent valables, à développer une IA qui soit conforme aux valeurs et à l'éthique européenne, protège la vie privée, lutte contre les discriminations ou les biais, accorde des droits aux personnes sur l'utilisation qui est faite des données qui les concernent. C'est à la construction de cet équilibre et pour préparer, en amont, l'adoption du règlement européen, qu'un événement comme celui d'aujourd'hui peut, j'en suis sûre certaine, contribuer très utilement.

Je vous remercie.