

Synthèse

Colloque

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

Vendredi 10 février 2023 de 9h00 à 15h45
Conseil d'État



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

Réussir l'équilibre entre protection des personnes et soutien à la recherche en santé

Dire que l'Intelligence Artificielle (IA) et les mégadonnées font aujourd'hui partie de nos vies quotidiennes est devenu un poncif ; communication, média, transports, banques, assurances, commerce, etc. sont autant de domaines où le recours à l'IA tend à devenir la norme avec des usages pour lesquels un équilibre entre protection des individus et innovation semble avoir été majoritairement trouvé.

Pour autant, un domaine de l'application de l'IA et des mégadonnées constitue un défi permanent aux acteurs publics et privés : celui de la santé. Il s'agit en la matière de trouver un équilibre entre protection des données autour de ce que l'individu a de plus intime et de plus précieux, sa santé, et le besoin légitime de la société de disposer de traitements et de diagnostics innovants pour répondre à des besoins médicaux insuffisamment ou mal couverts.

La collecte et l'exploitation des mégadonnées constituent aujourd'hui de très puissants leviers pour la recherche en santé mais viennent questionner les cadres réglementaires et d'organisation actuels de la recherche en France. Incontournables exigences de protection, mission d'expérimentation et d'innovation de la recherche publique à court et moyen terme, bénéfice sociétal à long terme, voici le triptyque au sein duquel les régulateurs et les législateurs doivent trouver un équilibre.

Pour mener la réflexion sur le cadre et l'organisation qui permettront de parvenir à cet équilibre, trois acteurs incontournables de l'utilisation de l'IA et des mégadonnées pour la recherche en santé que sont le Conseil d'Etat, la Commission nationale de l'informatique et des libertés (CNIL) et l'Alliance IHU France ont réuni leurs expertises avec l'organisation du colloque "IA et mégadonnées : comment vont-elles révolutionner la recherche et la pratique médicales de demain ?" en février 2023. Au cours de ce colloque, les intervenants ont présenté les perspectives des applications de l'IA autour de trois tables-rondes thématiques, qui ont abordé tour à tour les enjeux éthiques, juridiques et sociétaux de l'usage de l'IA dans la recherche en santé.

L'IA en Santé, une réalité quotidienne dans les laboratoires de recherche, des perspectives croissante pour les malades et les professionnels de santé

Dans les domaines de la pratique et de la recherche médicales, l'IA trouve d'ores et déjà de nombreux exemples d'applications concrètes mises au service du patient et des professionnels de santé : aide à la pose d'un diagnostic plus personnalisé, assistance dans les interventions chirurgicales, identification d'anomalies dans le parcours de soins, modélisation d'un jumeau numérique pour prescrire aux patients des traitements personnalisés, organisation des parcours de soins, etc. Au-delà d'approches personnalisées, le recours à l'IA pourra contribuer à répondre aux problématiques que rencontre actuellement le système de santé français. Ses capacités d'analyse, de modélisation et de suivi de données en temps réel en font une solution pour répondre aux disparités d'accès aux soins dans les territoires, et optimiser le fonctionnement et l'allocation des ressources dans le système de santé.

Ces exemples n'en sont encore qu'à leur début et les perspectives d'application qu'ouvrent des équipes multidisciplinaires de recherche foisonnent. Les acteurs du soin et de la recherche collaborent toujours plus avec des experts des sciences et technologies du numérique pour développer de façon ambitieuse leur exploitation des données de santé et leur usage de l'IA, comme le démontre le besoin croissant de bioinformaticiens dans les grands instituts de recherche. Ses centres d'excellence, son expertise médicale reconnue, ses experts, son système national de données de santé et son approche multidisciplinaire de la recherche clinique sont autant d'atouts qui peuvent permettre à la France de conjuguer recherche en santé et IA afin de pérenniser sa tradition d'excellence médicale et de la positionner en moteur européen de la recherche en santé.

L'encadrement des données de santé, entre consolidation et évolution

En Europe, le règlement général pour la protection des données (RGPD), le règlement sur les marchés numériques (DMA), le règlement sur les services numériques (DSA) et le règlement sur la gouvernance des données (DGA) posent une première base centrale pour l'exploitation, le partage et la gouvernance des données de santé. Ce cadre général vient s'ajouter au cadre propre à celui de la recherche en santé encadrant la recherche clinique. Il s'agit donc pour les chercheurs de maîtriser l'application de ces deux régimes. L'exercice est d'autant moins facile que la masse de données utilisées par la recherche encourage à collaborer à l'échelle européenne malgré la fragmentation des réglementations. Les acteurs de la recherche plébiscitent ainsi un cadre réglementaire transparent et lisible qu'ils peuvent s'approprier aisément.

Au-delà de la poursuite de la pédagogie et de l'accompagnement, le cadre de l'IA en santé ne doit pas pour autant faire croire que le travail de régulation des données de santé et de leurs usages est achevé. Ainsi que le souligne Thierry Tuot dans le rapport du Conseil d'Etat "S'engager dans l'intelligence artificielle pour un meilleur service public", l'évolution du cadre réglementaire doit remédier aux obstacles juridiques et techniques rencontrés par les acteurs de la recherche en élaborant une régulation des systèmes d'IA uniforme, répondant d'une part aux besoins d'innovation et de performance publique, et d'autre part aux impératifs de protection des droits et des libertés de la personne.

Penser un écosystème des données en santé en France conjuguant confiance et développement d'innovation

La réflexion sur l'usage de l'IA dans la recherche en santé n'est pas uniquement une question juridique. Il s'agit avant tout d'une question éthique, mais également économique, avec un rôle important joué par des acteurs privés. La consolidation de toutes ces dimensions est le défi auquel l'écosystème de l'innovation en santé doit répondre, ce qui surajoute au caractère sensible de l'utilisation des données de santé par la recherche.

L'exhaustivité de la collecte des données s'impose dans la mesure où elle doit permettre d'aborder de futures nouvelles questions de santé. C'est ainsi la question de la pertinence et de la qualité des données collectées mais également de leur utilisation postérieure par une diversité d'acteurs qui se pose. Pour autant que la constitution de ces bases de données soit un enjeu capital pour le développement et la variation de futures innovations thérapeutiques et diagnostiques, elle interroge sur la place du patient et de l'utilisateur dont le consentement éclairé doit être garanti.

Quels sont les enseignements du colloque ?

Intelligence artificielle et données massives en santé : de la promesse technologique à la transformation des pratiques

Une rupture méthodologique pour la recherche biomédicale

L'un des premiers enseignements des échanges concerne la nature même de la transformation en cours.

Comme l'a souligné le Pr Stéphane Hatem, l'IA ne signifie pas seulement l'accès à davantage de données ou à des outils diagnostiques plus performants. Les réseaux neuronaux et les stratégies d'apprentissage non supervisé interrogent le modèle traditionnel de production des connaissances fondé sur la séquence observation, hypothèse, expérimentation et interprétation. La capacité des algorithmes à identifier des associations complexes au sein de données massives pourrait ainsi faire émerger de nouvelles connaissances sans que le raisonnement qui les produit soit immédiatement accessible au chercheur. Cette évolution conduit directement aux problématiques d'explicabilité, mais également de propriété intellectuelle et de responsabilité.

De l'aide au diagnostic à la médecine de précision

Les cas d'usage présentés montrent une extension progressive de l'IA à différentes étapes de la prise en charge.

En diagnostic, l'analyse algorithmique permet d'assister le professionnel, de prioriser les cas nécessitant une expertise et potentiellement de gagner du temps. Elle peut également contribuer à renforcer l'expertise médicale et à identifier de nouveaux biomarqueurs, ouvrant la voie à une médecine de précision.

Dans le champ thérapeutique, l'intégration de données complexes et de jumeaux numériques permet d'envisager une assistance croissante aux gestes médicaux et à leur personnalisation.

Les essais cliniques constituent également un terrain d'évolution majeur. Les discussions ont notamment évoqué le recours aux dispositifs connectés pour enrichir le suivi, la sélection algorithmique de patients, la simulation de patients virtuels ou encore l'apprentissage fédéré permettant de partager de l'information sans nécessairement centraliser l'intégralité des données.

La donnée : première infrastructure de l'IA en santé

La performance d'un modèle dépend directement de la donnée utilisée pour son développement et sa validation.

Plusieurs enjeux ont été identifiés : qualité des données, fiabilité, accessibilité, interopérabilité et surtout représentativité des populations utilisées pour entraîner les algorithmes.

Des performances excessivement favorables peuvent ainsi résulter de biais méthodologiques ou de populations d'apprentissage insuffisamment représentatives. La validation clinique et méthodologique constitue donc une condition essentielle avant toute généralisation d'un outil.

La structuration des entrepôts de données hospitaliers représente dans ce contexte un enjeu stratégique. Leur hétérogénéité peut limiter les échanges et les usages, alors même qu'ils constituent une source particulièrement riche de données médicales. Les échanges ont ainsi fait ressortir la nécessité de favoriser leur homogénéisation et leur capacité à communiquer.

Partager sans nécessairement déplacer les données

La question de l'accès aux données traverse l'ensemble des débats.

Les techniques d'apprentissage fédéré constituent une piste importante : le modèle peut être entraîné localement, puis les résultats de cet apprentissage être partagés, plutôt que de faire circuler directement l'ensemble des données individuelles.

Le développement de données synthétiques représente une autre approche. Les chercheurs peuvent générer des données artificielles ou des cohortes virtuelles afin d'augmenter les volumes disponibles pour l'apprentissage. Ces approches pourraient également, dans certaines situations, faciliter la circulation des données lorsqu'elles ne sont plus directement rattachées à une personne.

Ces solutions ne suppriment cependant pas les questions méthodologiques et réglementaires : leur pertinence dépend du cas d'usage, de leur mode de génération et de leur capacité à représenter correctement la réalité clinique.

Explicabilité : rendre le résultat médicalement exploitable

La « boîte noire » algorithmique a occupé une place importante dans les échanges.

Pour le professionnel, l'enjeu n'est pas nécessairement d'être capable de restituer le fonctionnement mathématique du modèle. L'explicabilité doit surtout permettre de rendre son résultat cliniquement interprétable et transmissible.

Le professionnel doit être en mesure de comprendre le signal produit, de l'intégrer aux autres informations disponibles et de l'expliquer au patient. Cette approche rapproche finalement l'IA d'autres technologies médicales complexes : le médecin n'a pas besoin de maîtriser intégralement la technologie interne d'un appareil pour interpréter correctement le résultat qu'il produit, à condition d'en connaître les indications, les performances et les limites.

Cela implique nécessairement un enjeu de formation des professionnels de santé.

Quelle responsabilité dans une décision assistée par l'IA ?

Plus l'algorithme intervient dans la prise de décision, plus la question de la responsabilité devient concrète.

L'exemple évoqué lors du colloque est révélateur : lorsqu'un dispositif connecté détecte un risque élevé d'arythmie et d'accident vasculaire cérébral, qui assume la décision thérapeutique qui pourrait en découler ? L'algorithme, son concepteur, le fabricant du dispositif ou le médecin qui interprète le résultat ? Cette question rappelle que le déploiement de l'IA ne peut être dissocié de son intégration dans une chaîne de décision médicale clairement définie.

L'enjeu de l'évaluation doit également être proportionné au risque. Les discussions ont ainsi évoqué l'intérêt d'une cartographie des risques permettant d'adapter les modalités d'évaluation en fonction de l'impact potentiel du système d'IA sur la prise en charge.

Trouver l'équilibre entre protection et innovation

Les données de santé étant juridiquement considérées comme sensibles, leur traitement bénéficie d'une protection renforcée. Mais les échanges ont largement insisté sur la nécessité de ne pas opposer protection des personnes et progrès médical.

Le cadre juridique doit protéger la vie privée, la confidentialité et les droits fondamentaux tout en laissant à la recherche la possibilité de développer des innovations d'intérêt public. Il doit également être capable de s'adapter à l'évolution rapide des technologies.

La CNIL a ainsi souligné, que certains obstacles attribués au RGPD peuvent relever d'une interprétation excessive des règles et que des solutions sont souvent possibles.

L'objectif affiché est d'accompagner l'innovation tout en garantissant sa conformité.

Le référentiel relatif aux entrepôts de données de santé constitue un exemple de cette approche : le respect d'un cadre standardisé peut permettre leur constitution selon une procédure simplifiée, tout en maintenant un niveau élevé de protection.

Au-delà de la technologie : construire un écosystème économique

Après les usages et le cadre réglementaire, la dernière dimension majeure du colloque concernait le passage à la valorisation de la donnée.

Une innovation peut être scientifiquement pertinente et juridiquement possible sans jamais parvenir jusqu'au patient. Son développement suppose un accès aux données, des financements, des infrastructures, des compétences et des partenariats adaptés.

La question du modèle économique de la donnée de santé a ainsi été discutée, notamment au regard des différences entre modèles européens et extra-européens. En France, la construction d'un écosystème économique ne peut se réduire à la valorisation patrimoniale des données individuelles : la création de valeur doit également être recherchée dans les modèles, les services, les infrastructures, la recherche et les usages développés à partir de ces données.

Cette réflexion impose une coopération étroite entre acteurs publics et privés, producteurs de données, chercheurs, établissements de santé, industriels et financeurs.

Quelle responsabilité dans une décision assistée par l'IA ?

Le principal enseignement du colloque est probablement que le défi de l'IA en santé n'est plus uniquement technologique.

Les algorithmes existent et leurs applications médicales se multiplient. La prochaine étape consiste à créer les conditions permettant de transformer ces innovations en bénéfices démontrés pour les patients.

Cela suppose simultanément :

- Des données de qualité, représentatives et interopérables
- Une validation méthodologique et clinique rigoureuse
- Un cadre juridique protecteur mais compatible avec la recherche
- Une évaluation proportionnée aux risques
- Des résultats interprétables par les professionnels
- La formation des équipes médicales et scientifiques
- Une clarification progressive des responsabilités
- Des infrastructures et des modèles de financement permettant de valoriser la donnée de santé et permettre sa transformation en avancées concrètes pour les malades.

Et une coopération durable entre patients, soignants, chercheurs, institutions publiques et acteurs privés

L'enjeu n'est donc plus de savoir si l'intelligence artificielle prendra une place dans la médecine de demain, mais dans quelles conditions scientifiques, médicales, éthiques et économiques nous souhaitons organiser cette transformation pour continuer à jouer un rôle de premier ordre dans la recherche médicale au niveau international.

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

8h30 : accueil

9h00 : plénière d'ouverture

- Didier-Roland Tabuteau - Vice-Président du Conseil d'État
- Marie-Laure Denis - Présidente de la CNIL
- Stéphane Hatem - Directeur général de l'IHU ICAN

9h35 : présentation de l'étude du Conseil d'État sur l'IA et l'action publique

- Thierry Tuot - Président-adjoint de la section de l'intérieur du Conseil d'État

10h00 : l'usage de l'IA en recherche et dans la pratique médicale

- Modérateur : Gérard Biau - Directeur général de SCAI
- Stéphanie Combes - Directrice du Health Data Hub
- Dominique Le Guludec - Présidente de la HAS
- Pierre Jaïs - Directeur général de l'IHU Liryc
- Valérie Paradis - PUPH à l'AP-HP

11h15 : quel cadre réglementaire pour les données de santé ?

- Modérateur : Jean Lessi - Rapporteur à la section Sociale du Conseil d'État
- Antonios Bouchagiar, Membre du Service Juridique de la Commission Européenne
- David Gruson - Fondateur d'Ethik-IA
- Stéphane Hatem - Directeur général de l'IHU ICAN
- Valérie Peugeot - Commissaire responsable des données de santé à la CNIL

12h30 - Déjeuner libre

14h00 : quel écosystème économique pour les données de santé en France ?

- Modérateur : Louis Dutheillet de Lamothe - Secrétaire général de la CNIL
- Jérôme Chevallier - Directeur Stratégie & Opérations de Owkin
- Christian Deleuze - Directeur général délégué à l'Innovation Sanofi, Président de Medice Paris et Président du Healthcare Data Institute
- Sébastien Massart - Directeur de la stratégie chez Dassault Systèmes
- Isabelle Ryl - Directrice de PRAIRIE (PaRis Artificial Intelligence Research InstitutE)

15h10 : plénière de clôture

- Jean-Noël Barrot - Ministre délégué chargé de la Transition numérique et des Télécommunications
- Thierry Tuot - Président-adjoint de la section de l'intérieur du Conseil d'État

