Le dossier – L'intelligence artificielle et le cardiologue

De quelques problèmes posés par l'utilisation de l'intelligence artificielle en médecine

Il faut toujours prendre le maximum de risques avec le maximum de précautions.
Rudyard Kipling

L'arrivée massive de l'intelligence artificielle (IA) dans nos sociétés et dans la pratique médicale pose de nombreuses questions, allant des plus triviales aux plus complexes, des plus philosophiques aux plus réglementaires car, notamment, son objectif principal est de reproduire l'intelligence humaine. Certains y verront une aide, d'autres une substitution et une prise de contrôle. Dans cet article de conclusion de ce dossier sur l'IA seront abordés quelques-uns des problèmes qu'elle pose et ce, au-delà du risque d'erreurs qu'elle peut commettre dans les résultats qu'elle fournit.

➔ **F. DIÉVART**
DUNKERQUE.

Entre mythe et science, entre craintes ancestrales et promesses technologiques

Pour paraphraser un article de Pascal Lardellier et Emmanuel Carré paru le 5 janvier 2025 sur le site TheConversation.fr, la question de l'intelligence artificielle fait écho à des **débats millénaires sur notre rapport aux outils cognitifs**, questionnements de nature anthropologique et philosophique.

Faut-il verser dans un technosolutionnisme béat ou dans une technophobie primitive ? Le premier terme fait référence à un mode de pensée que l'on pourrait résumer comme suit : "Il n'y a pas de problème, il n'y a que des solutions", et ces solutions seront apportées par le développement technologique. Ce technosolutionnisme est conforté par l'augmentation de l'espérance de vie et l'amélioration des conditions de vie survenues depuis le début de l'ère industrielle : accès à une eau potable, accès aux soins, diminution de la pollution des villes… faisant qu'aujourd'hui les conditions de vie d'un Français sont meilleures que celles de Louis XIV.

Les tenants de la technophobie diront que ces progrès ont des contreparties, tels que le réchauffement climatique, l'épuisement des ressources de la planète, la perte du bon sens et de la vie de famille… Il est vrai que travailler aux champs 10 heures par jour avec les enfants et ne pas savoir si l'on pourra passer l'hiver avait de quoi faciliter la vie de famille avant l'ère industrielle, surtout si l'on était susceptible d'être enrôlé à tout moment dans l'armée de "ce bon Prince qui nous gouverne".

Même s'il peut en avoir l'apparence, l'article de ce dossier consacré à la consultation de cardiologie à l'ère de l'IA ne verse pas dans le technosolutionnisme béat mais a pour objectif de montrer une évolution quasi-inéluctable du parcours de soins, à quelques variantes près, et donc la nécessité qu'il y a de l'anticiper et de la considérer comme devant libérer du temps médical et de parole avec les patients. Certes, cette nouvelle donne créera des problèmes mais seront-ils supérieurs à ceux qu'elle va résoudre ? Seront-ils insurmontables ? Pour le dire autrement, qui, en France aujourd'hui, au nom des risques de la technique, aurait envie de mourir à 25 ans, des complications d'un rhumatisme articulaire aigu après avoir souffert toute sa vie de douleurs dues à des caries dentaires ou d'une dyspnée majeure ?

Un problème potentiel lié à l'IA est toutefois ce qui est constaté par certains médecins : la numérisation de la vie sociale (réseaux sociaux, courriels, visioconférences, SMS, alertes numériques induites par la télésurveillance, demandes incessantes de télé-expertise…) augmente la charge mentale avec son risque de surmenage professionnel (*burn-out*). En augmentant le flux des données numériques, l'utilisation de l'IA devra conduire à savoir organiser

Le dossier – L'intelligence artificielle et le cardiologue

son emploi du temps afin de préserver le temps de parole libre avec les patients.

Au-delà de la notion de technosolutionnisme, le fait que l'IA puisse se substituer à la pensée humaine doit-il être vécu comme une menace pour notre espèce ? Pour continuer dans le fil de l'article de P. Lardellier et E. Carré, il semble judicieux de concevoir que "*loin de remplacer nos capacités cognitives, les outils numériques les modifient et peuvent les enrichir, exactement comme l'écriture a fini par devenir un pilier de notre développement intellectuel plutôt qu'un substitut à la mémoire*" et "*c'est ici qu'intervient la distinction fondamentale entre prothèse et orthèse dans notre approche de l'IA. Une prothèse remplace une fonction défaillante ou manquante, tandis qu'une orthèse soutient et améliore une fonction existante*".

Il semble que cette conception de l'IA soit celle à retenir, laissant de côté ce qui reste encore un mythe, la notion de la prise de pouvoir des machines sur l'homme… Ceci n'empêche pas de réfléchir sur quelques limites et dangers de l'IA.

L'automatisation des soins et son risque de déshumanisation

Un de ces risques est potentiellement celui d'une déshumanisation de la relation médecin-patient. L'IA, par l'automatisation des tâches médicales qu'elle va permettre, doit conduire à libérer du temps médical et non à remplacer le temps de dialogue en créant de la distance entre le médecin et le patient. Ce sera donc ici une affaire de personnes où certains s'épanouiront dans le soin et l'accompagnement, et d'autres laisseront la machine faire : "si vous voulez un antipyrétique tapez 1 ; si vous voulez un antibiotique tapez 2 ; si vous voulez une assistance vocale tapez 3… ".

Un autre risque concerne la décision médicale. Ici, il y a encore une grande incertitude quant à l'évolution de l'environnement médico-légal. Les données du problème sont les suivantes : la prise de décision (diagnostique et thérapeutique) repose sur le médecin qui a été formé pour cela et, dans chaque décision, il engage sa responsabilité médico-légale ; les outils à base d'IA pourraient améliorer et améliorent déjà la précision diagnostique, la qualité du suivi et l'adaptation du traitement dans de nombreuses situations cliniques ; ainsi, le médecin pourrait être tenté (par habitude ou excès de confiance notamment) de se fier aux résultats fournis par l'IA et à ses préconisations, et il serait même possible d'envisager un support légal faisant que si l'IA est plus précise que le médecin, ce dernier devrait alors prendre ses préconisations en compte.

On est ici confronté au problème d'arbitrage entre deux modes décisionnaires dont on ne sait comment il évoluera, ce d'autant plus que si l'on pense que le médecin doit être formé pour superviser les préconisations et résultats fournis par des IA, on peut aussi penser l'inverse, l'IA pourrait rectifier de façon adaptée le raisonnement du médecin.

Dans combien de temps ce débat émergera-t-il ? Nul ne le sait, mais il se posera et devra mettre en avant des principes, dont l'un pourrait être l'efficacité. Au-delà des débats philosophiques, le problème actuel est de comprendre certaines des limites de l'IA qui posent problème pour son utilisation médicale.

L'explicabilité

Un des problèmes majeurs posés par l'IA est l'explicabilité de ses résultats. L'explicabilité est la capacité à mettre en relation et à rendre compréhensibles les éléments pris en compte par le système d'IA pour la production d'un résultat. Dans un système expert, comme peut l'être un texte de recommandations, l'homme construit puis utilise des algorithmes : dans telle situation clinique, les données de la science justifient de faire tel bilan, de proposer tel traitement en première intention, en cas d'aggravation, de compléter le bilan par certains éléments et de proposer, éviter ou arrêter tel autre traitement.

Dans une IA reposant sur l'apprentissage profond, c'est l'IA elle-même, qui, à partir de modèles statistiques appliqués à la base de données qu'elle analyse, construit le chemin entre les éléments entrants et les éléments sortants avec des critères peu ou pas compréhensibles. Ainsi, par exemple, en analysant une base de données de milliers d'ECG et en ayant comme élément entrant le fait que parmi tous les patients auxquels ont été fait ces ECG, certains ont de la fibrillation atriale (FA) paroxystique, l'IA va pouvoir déceler des caractéristiques non visibles par l'homme dans les ECG des patients ayant une FA paroxystique. Elle peut ensuite conclure face à un nouvel ECG qui lui est soumis que le patient a de la FA paroxystique ou non et ce, sans que l'on sache quels critères elle a retenu pour ce faire. On ne peut expliquer sa démarche, tout au plus peut-on l'évaluer en vérifiant que sa prédiction est exacte et reproductible dans des populations différentes que celles lui ayant servi de base d'entrainement. Et ce, notamment si sa base de données ayant servi à l'analyse n'a pas comporté de biais de sélection, en d'autres termes, si elle a été représentative de l'ensemble de la population.

Il est à noter que c'est au terme d'une telle évaluation en population que cette IA devrait ou non être approuvée et utilisée. Et cela parait indispensable car lorsqu'on analyse des milliers voire des milliards de données, il est toujours possible, sinon probable de trouver des **corrélations sans fondement ou des corrélations parasites,** c'est-à-dire des corrélations que rien n'explique mais qui sont temporairement validées par de grandes séries statistiques. Ainsi, aux USA, il existe un site (*spurious correlation*) analysant les données statistiques du Web et ayant permis de constater qu'il existe

plusieurs corrélations entre diverses données, dans une période donnée, parfois pendant plusieurs années sans aucune explication logique autre que le hasard. Il en est ainsi, par exemple, de l'évolution des dépenses publiques américaines dans la recherche scientifique astronomique et technologique et du taux de suicide par pendaison, étouffement et étranglement : ces deux variables ont évolué de façon strictement parallèle pendant 20 ans (faut-il alors réduire les dépenses en recherche pour diminuer le taux de suicide par pendaison ?). On trouvera également une corrélation entre les personnes se noyant dans des piscines par rapport au nombre d'apparitions de Nicolas Cage à l'écran (ici, est-on bien certain qu'il s'agisse d'un hasard ?), de la consommation de mozzarella par tête par rapport au nombre de doctorants en ingénierie civile (les Italiens diront certainement que ce n'est pas le hasard), le nombre de morts liés à une chute depuis un bateau de pêche par rapport au nombre de mariages dans le Kentucky (est-ce un effet de la législation sur les rames aux USA ?)…

Vers l'IA explicable

Outre le débat concernant la différence entre causalité et corrélation, le problème réellement posé est que l'homme souhaite comprendre le mécanisme conduisant à proposer un résultat plutôt qu'un autre lorsque des données sont analysées. Face à ce problème, il y a deux grandes écoles :
– une première considère que le processus même des réseaux neuronaux ne pourra pas permettre (du moins à court et moyen terme) d'expliquer les modalités de production d'un résultat et qu'il faut donc utiliser l'IA telle qu'elle est actuellement, sans rendre encore plus complexe sa mise à disposition ;
– une deuxième considère qu'il faut, pour des raisons scientifiques et éthiques, évoluer vers une IA explicable (dénommée en anglais XAI) et ce, malgré les difficultés rencontrées.

Les lignes qui suivent sont adaptées de données provenant de la société IBM qui fait partie des sociétés travaillant sur l'IA explicable :
– à mesure que l'IA se perfectionne, un défi est de comprendre comment l'algorithme est parvenu à chaque résultat car l'ensemble du processus de calcul est retranscrit dans ce qu'on appelle une "boîte noire", créée directement à partir des données et qui est impossible à interpréter, même par les ingénieurs ou *data scientists* qui créent l'algorithme ;
– il y a de nombreux avantages à comprendre comment un système basé sur l'IA a abouti à un résultat spécifique. L'explicabilité peut ainsi aider les développeurs à **vérifier que le système fonctionne comme prévu**, permettre de **répondre aux normes réglementaires**, ou encore **permettre aux personnes concernées par une décision de contester ou de modifier ce résultat ;**
– l'IA explicable ou XAI est un ensemble de processus et de méthodes qui permettent aux utilisateurs humains de comprendre et de faire confiance aux résultats créés par les algorithmes de *machine learning* ;
– l'IA explicable est utilisée pour décrire un modèle d'IA, son effet attendu et ses biais potentiels ;
– elle permet de caractériser la précision, l'équité, la transparence et les résultats du modèle dans la prise de décision alimentée par l'IA. L'IA explicable est cruciale pour une organisation, car elle permet d'instaurer un climat de confiance lors de la mise en production de modèles d'IA. Cette explicabilité aide également une organisation à adopter une approche responsable du développement de l'IA ;
– une organisation ne doit pas suivre l'IA sans esprit critique. Il est impératif qu'elle comprenne parfaitement son processus décisionnel, ce que permet un modèle de contrôle et de responsabilité de l'IA.

Au-delà du problème des corrélations parasites, ces éléments soulignent certains des défauts potentiels des résultats fournis par une IA dans le domaine médical :
– en analysant des corpus de données fournis par des recherches conduites par l'homme et chez l'homme, lesdits résultats en reproduisent les biais, notamment les biais de sélection (populations masculines, jeunes et caucasiennes surreprésentées dans les registres et essais cliniques, patients hospitalisés différents de ceux de la population générale), mais aussi les préjugés fondés sur la race, le sexe, l'âge ou le lieu de résidence) ;
– les performances des modèles peuvent dériver ou se dégrader, car les données de production peuvent différer au fil du temps justifiant de surveiller et de gérer en continu les modèles afin de promouvoir l'explicabilité de l'IA.

Comment fonctionne l'IA explicable

Le défi est de franchir le pas séparant interprétation et explication des processus d'IA. L'interprétabilité est la mesure dans laquelle un observateur peut comprendre la cause d'une décision. Il s'agit du taux de réussite que les humains peuvent prédire pour les résultats d'une IA.

L'IA parvient à un résultat à l'aide d'un algorithme de *machine learning* (ML), mais les architectes des systèmes d'IA ne comprennent pas entièrement comment celui-ci est parvenu à ce résultat. Il est donc difficile d'en vérifier l'exactitude, ce qui entraîne une perte de contrôle, de responsabilité et d'auditabilité. La XAI utilise des techniques et des méthodes spécifiques pour garantir que chaque décision prise au cours du processus de ML peut être retracée et expliquée.

Les techniques de XAI reposent sur trois axes principaux. La précision et la traçabilité des prédictions qui répondent aux exigences technologiques, et la compréhension des décisions qui répond aux besoins humains.

Le dossier – L'intelligence artificielle et le cardiologue

La précision est un élément clé de la réussite de l'utilisation de l'IA dans les opérations quotidiennes. En effectuant des simulations et en comparant les résultats de la XAI à ceux du jeu de données d'entraînement, il est possible de déterminer la précision de chaque prédiction. La technique la plus utilisée à cet effet est LIME (pour *Local Interpretable Model-Agnostic Explanations*), qui explique la prédiction des classificateurs par l'algorithme de *machine learning*.

La traçabilité est une autre technique clé de la XAI. Cela se fait, par exemple, en limitant la manière dont les décisions peuvent être prises et en établissant un champ d'application plus étroit pour les règles et la fonctionnalité du ML. Parmi les techniques de traçabilité XAI, il y a DeepLIFT (pour *Deep learning Important FeaTures*), qui compare toute activation de neurone à son neurone de référence et trace le lien entre chaque neurone activé, et même les dépendances entre eux.

La compréhension des décisions est le facteur humain. Comme de nombreuses personnes se méfient de l'IA, pour l'utiliser efficacement, elles doivent apprendre à lui faire confiance en comprenant comment et pour quelles raisons elle prend ses décisions.

De quelques avantages de l'IA explicable

L'IA explicable permet de revoir et d'améliorer les performances des modèles tout en aidant les parties prenantes à comprendre les comportements de l'IA. Une étude comportementale des modèles par le suivi des informations sur l'équité, la qualité, la dérive et l'état du déploiement est indispensable pour développer l'IA à grande échelle.

L'évaluation continue des modèles permet de comparer leurs prédictions, de quantifier les risques et d'optimiser leurs performances : un processus qui peut être accéléré en affichant des valeurs positives et négatives sur les comportements des modèles avec les données qui génèrent l'explication. En rendant les modèles d'IA explicables et transparents, il devient possible de gérer la réglementation, la conformité, les risques et les autres exigences, de minimiser les frais liés aux inspections manuelles et aux erreurs coûteuses, et d'atténuer les risques de biais involontaires.

En conclusion

En médecine, l'explicabilité des résultats d'une IA parait essentielle au point que certains souhaitent qu'elle fasse partie du cahier des charges avant approbation par une agence réglementaire et donc avant son utilisation. Si les avantages sont évidents, les obstacles sont nombreux. Outre la longueur et la complexité du processus, cela justifierait d'avoir accès aux bases de données ayant permis l'entrainement de l'IA et donc d'analyser si les sources utilisées ont respecté les droits des producteurs et des propriétaires de ces sources.

Les sources, leur protection et le droit

Plus le nombre de données analysées est grand, plus grande sera la précision d'une mesure : c'est un principe de base. Par exemple, si une chaine de production commet une erreur en moyenne sur 1 % des produits qu'elle assemble, vous améliorez la connaissance de la précision de ce taux en augmentant le nombre de produits que vous allez examiner. Si vous examinez 10 produits, votre marge d'erreur sera telle qu'aucune conclusion ne sera possible, si vous en examinez 100 vous pouvez tout à fait avoir 0 produit défectueux comme vous pouvez en avoir 3 à 4 ; si vous en examinez 10 000, il est probable que vous soyez très proche du taux réel de 1 % de produits défectueux. Il en est de même pour créer une inférence statistique : plus l'IA peut analyser de données, plus précis sera le résultat qu'elle proposera.

Ainsi, ChatGPT a analysé plusieurs milliards de données et il est dit que cela a été possible par un véritable pillage des bases de données numériques du Web, indépendamment des droits afférents à ces données puisqu'il a analysé, au-delà des bases de données publiques fournies par Common Crawl, des articles de presse ou des livres, voire selon certains auteurs, le contenu de courriels et de fils WhatsApp pour adapter ses réponses à des profils de langage plus familiers. Cela lui a d'ailleurs valu une interdiction transitoire de fonctionnement en Italie (en 2023) au motif qu'OpenAI, sa société mère "a traité les données personnelles d'utilisateurs pour former ChatGPT sans disposer d'une base juridique adéquate", enfreignant "le principe de transparence et les obligations d'information envers les utilisateurs qui en découlent".

En médecine, le problème est le même : pour qu'une IA puisse s'entrainer, il lui faut disposer de données et évidemment de données médicales, donc personnelles et sensibles. Comment obtenir un volume suffisant de données adaptées ? Selon quelle réglementation et avec quelle protection pour les personnes dont les données sont exploitées ? La recherche médicale, de quelque nature qu'elle soit est réglementée et la vente de données par les médecins est interdite en France, la pseudonymisation des données étant fragile.

De plus, les réglementations française et européenne donnent aux personnes fournissant des données à une IA (voire à tout médecin conservant des données relatives à un patient) le droit à l'information qui implique une politique de confidentialité et d'information comprenant :
– ce qui est collecté et par qui ;
– pour combien de temps ;
– en vertu de quel fondement juridique ;
– pour quel objectif et comment ces données sont sécurisées.

Le dossier – L'intelligence artificielle et le cardiologue

En Europe, et donc en France, le RGPD prévoit plusieurs droits à disposition des personnes :
– le droit d'accéder à leurs données personnelles ;
– le droit de demander une rectification de ces données ;
– le droit de s'opposer à l'utilisation qui en est faite ;
– le droit de demander la suppression de leurs données personnelles ;
– et le droit d'obtenir une copie de ces données.

L'Europe a par ailleurs émis un règlement dénommé IA Act qui encadre les modalités de développement et d'utilisation de l'IA (***encadré I***).

À tout moment donc, le médecin collectant des données de santé doit mettre à disposition les moyens nécessaires pour faciliter l'exercice de ces droits. Plus encore, à toute étape du cycle de vie d'une donnée, celui qui les collecte et les conserve doit être vigilant sur le mode de collecte, sur leur utilisation et la manière dont elles sont sécurisées :
– ainsi, les données personnelles doivent être collectées en vertu d'une autorisation. Le RGPD prévoit 6 hypothèses limitatives autorisant la collecte puis l'utilisation des données. Appelé "base légale", ce principe signifie que **les données personnelles ne peuvent être traitées que sur un fondement juridique.** L'article 6 RGPD en énumère 6 : le consentement de la personne, l'exécution d'un contrat, le respect d'une obligation légale, l'intérêt légitime de l'organisme, l'intérêt public ou l'intérêt vital ;
– l'article 5 du RGPD, qui énumère les principes fondateurs du RGPD prévoit que protéger les données personnelles implique notamment de conserver l'intégrité des informations utilisées et de ne pas altérer leur exactitude ;
– enfin, tout organisme doit, en toute circonstance, être en mesure d'assurer la sécurité des données contre toute violation, fuite, altération ou destruction en prévoyant les mesures de sécurité adéquates (article 32 RGPD).

L'IA Act européen

Les États de l'Union européenne (UE) ont approuvé un règlement européen sur l'IA le 21 mai 2024. Le Conseil de l'Europe a adopté le 17 mai 2024 un traité international visant à garantir une IA respectueuse des droits fondamentaux. Les États membres de l'UE ont validé la proposition de règlement sur l'IA à l'unanimité.

Il s'agit d'une législation inédite au niveau mondial pour réguler l'IA.

Les auteurs de ces textes ont tenté de trouver un équilibre entre la protection contre les dangers de l'IA et la volonté de ne pas prendre de retard dans son développement en Europe. Ainsi, cette approche a voulu prendre en compte les résultats bénéfiques sur les plans sociaux et environnementaux que peut apporter l'IA, mais aussi les nouveaux risques ou les conséquences négatives que peut engendrer cette technologie.

Son objectif a été :

- D'établir des directives claires et précises à l'attention des développeurs et des utilisateurs d'IA.
- De réduire les contraintes administratives et financières imposées aux entreprises, notamment les petites et moyennes entreprises (PME).
- De favoriser le développement d'une IA de confiance, garantissant les droits fondamentaux, la sécurité et les principes éthiques, tout en encourageant et en renforçant l'investissement et l'innovation de l'IA dans l'ensemble de l'UE.
- De veiller à ce que les systèmes d'IA mis sur le marché soient sûrs et respectent la législation en vigueur en matière de droits fondamentaux, les valeurs de l'UE, l'État de droit et la durabilité environnementale.
- De garantir la sécurité juridique afin de faciliter les investissements et l'innovation dans le domaine de l'IA.
- De renforcer la gouvernance et l'application effective de la législation existante en matière d'exigences de sécurité applicables aux systèmes d'IA et de droits fondamentaux.
- De faciliter le développement d'un marché unique pour des applications d'IA légales et sûres, et empêcher la fragmentation du marché.

Ce règlement établit :

- L'interdiction de certaines pratiques.
- Des exigences spécifiques applicables aux systèmes d'IA à haut risque.
- Des règles harmonisées en matière de transparence applicable :
  – aux systèmes d'IA destinés à interagir avec des personnes,
  – aux systèmes de reconnaissance des émotions et de catégorisation biométrique,
  – aux systèmes d'IA générative utilisés pour générer ou manipuler des images ou des contenus audio ou vidéo.

Cette proposition complète le droit existant en matière de non-discrimination en prévoyant des exigences qui visent à réduire au minimum le risque de discrimination algorithmique, assorties d'obligations concernant les essais, la gestion des risques, la documentation et le contrôle humain tout au long du cycle de vie des systèmes d'IA.

Encadré I (1re partie) : L'IA ACT européen.

Les produits et droits

Quand on utilise une IA, on utilise un service connecté à un serveur (*cloud*) qui reçoit les données pour les analyser. Quel est le risque qu'elles soient piratées ? Qu'elles puissent être fournies à une autorité qui en fait la demande ? Qui est propriétaire des données ? Qui est propriétaire des résultats de l'analyse (***encadré II***) ?

Ce qui est certain, c'est que le médecin qui utilise une IA doit être vigilant car il s'expose à de nombreuses fautes ou erreurs. Le patient doit être informé du traitement de ses données par une IA, les outils à base d'IA utilisés dans le domaine médical doivent répondre à un cahier des charges de protection des données, de confidentialité et de sécurité qu'il est prudent de vérifier avant l'utilisation.

Plus encore, toute utilisation de données personnelles d'un patient dans une IA non médicale ne répondant pas à un cahier des charges précis en matière de sécurité des données expose à des poursuites : ainsi, si l'on veut demander à ChatGPT des précisions sur la prise en charge médicale d'une personne, la question ne doit pas permettre d'identifier cette personne. Plus encore, les politiques de confidentialité de ChatGPT, mais aussi de l'ensemble des IA génératives ne garantissent pas les droits des personnes conformément au RGPD. Notamment, il n'existe aucune information concernant le fait que les conversations qu'ils entretiennent avec ChatGPT servent à alimenter et perfectionner l'algorithme et qu'elles sont ainsi toutes enregistrées. ChatGPT n'informe pas ses utilisateurs sur le fondement légal qui l'autoriserait à collecter et conserver toutes les informations qui lui sont transmises dans le cadre de son utilisation et il n'existe aucun fondement juridique tel que le consentement, le contrat ou l'intérêt légitime autorisant la collecte de ces données en toute sécurité.

Ses utilisateurs ne sont pas informés des droits qu'ils peuvent exercer (accès, suppression et opposition au traitement qui sont les plus importants) et n'ont aucun moyen de les exercer. Et ce n'est pas tout, car comme l'écrit une avocate en droit public (https://www.leto.legal/guides/intelligence-artificielle-chatgpt-et-rgpd) : "*ChatGPT enregistre l'intégralité des informations qui sont transmises à l'IA. Une attaque cybercriminelle engendrerait un risque incalculable pour le respect de la vie privée de ses utilisateurs. Or, aucun élément ne permet de s'assurer des mesures de sécurité mises en place par OpenAI pour réduire ce risque… OpenAI est une entreprise américaine dont le droit auto-*

L'IA Act européen

Les pratiques suivantes en matière d'IA sont interdites :

- Les systèmes d'IA ayant recours à des techniques subliminales au-dessous du seuil de conscience d'une personne pour altérer substantiellement son comportement de manière à causer un préjudice physique ou psychologique (manipulation du comportement humain pour contourner le libre arbitre).

- Les systèmes d'IA exploitant les éventuelles vulnérabilités dues à l'âge ou au handicap d'un individu pour altérer substantiellement son comportement et de manière à causer un préjudice physique ou psychologique.

- Les systèmes d'IA destinés à évaluer ou à établir un classement de la fiabilité de personnes en fonction de leur comportement social ou de caractéristiques personnelles et pouvant entraîner un traitement préjudiciable de personnes, dans certains contextes, injustifié ou disproportionné. L'accord trouvé entre le Parlement et les États membres précise l'interdiction des systèmes de catégorisation biométrique utilisant des caractéristiques sensibles (opinions politiques, religieuses, philosophiques, orientation sexuelle…) et la notation sociale basée sur le comportement social ou les caractéristiques personnelles.

- Les systèmes d'IA pour mener des évaluations des risques des personnes physiques visant à évaluer ou à prédire le risque qu'une personne physique commette une infraction pénale, uniquement sur la base du profilage d'une personne physique ou de l'évaluation de ses traits de personnalité ou caractéristiques. Cette interdiction ne s'applique pas aux systèmes d'IA utilisés pour étayer l'évaluation humaine de l'implication d'une personne dans une activité criminelle.

- Les systèmes d'IA créant ou développant des bases de données de reconnaissance faciale par le moissonnage non ciblé d'images faciales provenant d'internet ou de la vidéosurveillance.

- La reconnaissance des émotions sur le lieu de travail et les établissements d'enseignement, sauf pour des raisons médicales ou de sécurité.

- Les systèmes d'identification biométrique à distance en temps réel dans des espaces accessibles au public à des fins répressives, sauf dans les cas suivants : recherche ciblée de victimes potentielles spécifiques de la criminalité (enfants disparus, traite, exploitation sexuelle), prévention d'une menace spécifique, substantielle et imminente pour la vie ou la sécurité des personnes ou la prévention d'une attaque terroriste, identification, localisation ou poursuite à l'encontre des auteurs ou des suspects de certaines infractions pénales punissables d'une peine d'une durée maximale d'au moins quatre ans.

L'utilisation de systèmes d'identification biométriques à distance en temps réel doit :
– tenir compte de la situation donnant lieu au recours au système et de la gravité ou de l'ampleur du préjudice en l'absence de son utilisation ;
– tenir compte des conséquences sur les droits et libertés de toutes les personnes concernées (gravité, probabilité, ampleur) ;
– être subordonnée à une autorisation préalable octroyée par une autorité judiciaire ou administrative compétente.

Encadré I (2e partie) : L'IA ACT européen.

Le dossier – L'intelligence artificielle et le cardiologue

IA générative et protection des données

La plupart des modèles d'IA générative sont hébergés sur des serveurs américains et, en vertu du *Patriot Act* et du *Cloud Act*, les données envoyées peuvent être récupérées par les autorités américaines.

Les données fournies aux IA génératives sont réutilisées pour améliorer les modèles, donnant la possibilité de retrouver ces données lors de futures interrogations. Cela peut représenter un risque en matière de sécurité et de confidentialité des données.

Il existe quelques solutions d'hébergement avec des espaces dédiés et fermés, ou encore des alternatives d'IA génératives *open source* pouvant être installées sur des serveurs locaux.

Encadré II : IA générative et protection des données.

rise les renseignements à avoir accès à n'importe quelle information détenue par une entreprise immatriculée aux États-Unis ce qui contrevient très largement au RGPD".

En conclusion

Si l'utilisation de l'IA est promue comme devant faire profondément évoluer la pratique de la médecine, elle pose encore de nombreux problèmes de divers ordres et notamment philosophiques car elle a pour objectif de simuler les fonctions cognitives de l'homme ; juridiques car elle nécessite l'exploitation de données personnelles et scientifiques, et nécessite d'être précise, reproductible et donc évaluée afin de connaître son apport réel à la pratique médicale.

Plus encore, malgré sa croissance qui parait exponentielle, elle pose le problème de son évolution réelle lorsque l'on sait que son coût d'exploitation en termes d'énergie, de refroidissement et de puissance de calcul la rend non viable, tant sur le plan commercial que dans le contexte de la crise climatique.

L'auteur a déclaré les liens d'intérêts suivants : honoraires pour conférences ou conseils ou défraiements pour congrès pour et par les laboratoires Alliance BMS-Pfizer, Amgen, Astra-Zeneca, Bayer, BMS, Boehringer-Ingelheim, Menarini, Novartis, Novo-Nordisk, Pfizer, Sanofi France, Servier.