Une indispensable réflexion éthique en amont de l'utilisation de l'IA dans le domaine de la santé

Nathanaël Laurent[a]
Formateur en méthode de travail, docteur en sciences biomédicales, doctorant en philosophie.

Federico Giorgi[a]
Doctorant

Mathieu Guillermin[b*]
Philosophe des sciences, maître de conférences

[a]Institut de recherche "Espace philosophique de Namur" (Esphin), Université de Namur, 61 rue de Bruxelles, 5000 Namur, Belgique

[b]UR Confluence : sciences et humanités (EA 1598), Lyon Catholic University, 10 place des Archives, 69002 Lyon, France

L'utilité des outils d'intelligence artificielle (IA) dans les domaines de la santé et de la médecine ne fait aucun doute. Il est néanmoins important de prendre en compte certaines difficultés : comment préserver la relation médecin-patient dans un environnement de plus en plus dirigé et contrôlé par des IA statisticiennes ? Comment préserver, au cœur de la relation de soins, les pratiques médicales qui se nourrissent des singularités et du non prédictible ? La médiation de la technique ne doit-elle pas être évaluée dans la pratique médicale et avec les patients, plutôt qu'imposée *a priori* comme s'il s'agissait de la seule possibilité d'action ?

© 2024 Elsevier Masson SAS. Tous droits réservés, y compris ceux relatifs à la fouille de textes et de données, à l'entraînement de l'intelligence artificielle et aux technologies similaires.

Mots-clés – éthique ; humain ; intelligence artificielle ; médecine ; technique

Ethical reflection is essential prior to the use of AI in healthcare. The usefulness of artificial intelligence (AI) tools in the fields of health and medicine is beyond doubt. However, it is important to take certain difficulties into account: how can we preserve the doctor-patient relationship in an environment increasingly directed and controlled by statistician AIs? How can we preserve, at the heart of the healthcare relationship, the medical practices that thrive on singularities and the unpredictable? Shouldn't the mediation of technology be evaluated in medical practice and with patients, rather than imposed a priori as if it were the only possible course of action?

© 2024 Elsevier Masson SAS. All rights reserved, including those for text and data mining, AI training, and similar technologies.

Keywords – artificial intelligence; ethics; human; medicine; technology

Dans le domaine de la santé, comme dans tout autre, les techniques d'intelligence artificielle (IA) sont de plus en plus nombreuses et prometteuses. Pour chaque tâche requérant de traiter avec précision de nombreuses informations, un outil d'IA semble désormais pouvoir intervenir. Plus rapides, plus efficaces, plus précises, mais surtout capables de jongler avec d'immenses quantités de données (*big data*) en tenant compte de toujours plus de paramètres, ces techniques informatiques ne cessent de susciter des espoirs. Fruits de la recherche et donc de l'intelligence humaine, les IA seraient comme un prolongement de cette dernière permettant d'apporter de nouvelles solutions à nos problèmes.

À lire certains acteurs de la santé, la voie de l'IA en santé et en médecine serait toute tracée ; comme s'il s'agissait d'une réponse quasi miraculeuse aux limites de nos systèmes de soins : *« D'une part, des plaies structurelles grandes ouvertes : erreurs de diagnostic, variabilité des situations, faillibilité des experts, grandes difficultés à transmettre les informations de la recherche vers les praticiens. D'autre part, une IA qui excelle à digérer des monceaux de littératures, à dénicher des corrélations rares, à analyser les images et autres données, toujours plus nombreuses, que produit la médecine. Et bien sûr, un secteur où les enjeux sont littéralement vitaux. »* [1]

On ne parle pas ici de l'idée fantaisiste de remplacer le médecin, mais de l'objectif bien plus concret et réaliste de l'augmenter : *« Le médecin automatique n'est certainement pas pour demain, et n'est certainement pas souhaitable, mais le médecin "augmenté", mieux outillé pour guérir, prévenir, analyser, décider, et en même temps pourvu d'empathie et doué pour le contact humain, a certainement de beaux jours à venir. »* [1]

À quoi pourrait ressembler demain cette médecine augmentée ? Personne ne le sait ! Qu'est-ce qui peut donc donner autant confiance

*Auteur correspondant.
Adresse e-mail : mguillermin@univ-catholyon.fr (M. Guillermin).

© 2024 Elsevier Masson SAS. Tous droits réservés, y compris ceux relatifs à la fouille de textes et de données, à l'entraînement de l'intelligence artificielle et aux technologies similaires.
http://dx.doi.org/10.1016/j.scad.2024.12.011

© 2025 Elsevier Masson SAS. Tous droits réservés. - Document téléchargé le 07/03/2025 par CH BELFORT MONTBELIARD - (97567). Il est interdit et illégal de diffuser ce document.

à ces acteurs de la santé ? Comment peuvent-ils être aussi certains que la médecine augmentée par l'IA telle que la médecine prédictive ou de précision est la médecine souhaitable de demain, et que les outils d'IA sont en quelque sorte les moyens attendus depuis toujours – comme s'ils étaient inscrits dès le départ dans le projet de la médecine ?

Un monde dominé par les IA

Ce monde de la médecine prédictive rendu possible par les nouvelles technologies de l'informatique est bel et bien déjà dans la tête de nombreux acteurs de la santé, comme en témoignent les propos de Kerrie Holley et Siupo Becker, auteurs de *AI-First Healthcare : AI Applications in the Business and Clinical Management of Health* : *« Les réalisations d'IA dominent notre monde, au travail, dans les loisirs et à la maison. Aujourd'hui, les prestataires de soins de santé et les patients vivent à une époque où de nombreuses prédictions d'états pathologiques se produisent avant que le patient ne tombe malade. La présence croissante d'objets intelligents crée de vastes opportunités dans le domaine des soins de santé. Comme l'explosion cambrienne il y a 500 millions d'années, un moment cambrien de l'intelligence des machines, des dispositifs fusionnés avec l'IA, est arrivé, avec d'immenses possibilités de transformation pour les soins de santé »* [2].

Pour ces auteurs et, à n'en pas douter bien d'autres, *« l'époque où un médecin déclarait : "D'après mon expérience, ce traitement fonctionne" est révolue »*. Selon eux l'arrivée des outils d'IA permet désormais aux médecins de dire : *« Mon expérience, ainsi que celle de centaines de milliers de patients, de collègues médecins et d'études cliniques, me donnent confiance pour poursuivre cette voie thérapeutique. »* [2]

S'il en va ainsi, la pratique médicale s'apprête à être profondément chamboulée. Le nouveau partenariat techno-humain qui est en train de se mettre en place modifiera en effet considérablement la relation médicale. L'IA est vouée à y prendre place pour devenir un partenaire à part entière. *« Il en résulte un partenariat entre l'homme et l'IA, dans lequel l'IA exécute rapidement, efficacement et avec précision la tâche qui lui est confiée, et l'homme effectue son évaluation plus complexe et nuancée, en tenant compte des facteurs humains supplémentaires qui peuvent avoir un impact sur la solution pour le patient. »* [2]

Comment s'assurer que dans ce triangle IA-médecin-patient, le patient reste bien le principal bénéficiaire de la relation de soins ? Le domaine de la santé peut-il s'assurer de ne pas succomber au "capitalisme de surveillance" étudié en détail par Shoshana Zuboff [3] ? Qui contrôlera les informations et données personnelles générées par les capteurs, enregistreurs et autres détecteurs qui envahissent notre environnement et qui s'avèrent nécessaires pour rendre possible une médecine personnalisée et de précision ? À qui profiteront ces immenses banques de données médicales, alors que l'on observe déjà la concentration des IA les plus avancées entre les mains d'un nombre restreint d'oligarques numériques, leur conférant un pouvoir d'influence considérable sur certains comportements humains par le biais de l'analyse prédictive et de la recommandation de contenus ? Qui peut garantir aujourd'hui que les patients ne seront pas réduits à des objets numériques, à des données se perdant dans le dédale de gigantesques calculs de probabilités au sein d'IA de type *deep learning* ?

De son côté, Zuboff affirme que le capitalisme de surveillance sape la démocratie en privatisant l'information et en manipulant les préférences, c'est-à-dire globalement en affaiblissant la capacité des citoyens à prendre des décisions éclairées. Dans un esprit similaire, nous pouvons nous interroger : la médecine prédictive (et donc de surveillance) ne sape-t-elle pas la pratique médicale en déshumanisant la relation médecin-patient ? En cherchant avant tout à réduire le risque d'erreur, et en visant la performance maximale, cette médecine ne risque-t-elle pas de passer à côté de la vocation première du soin et de la médecine : une pratique centrée sur la rencontre avec le vivant humain dans toute sa complexité ?

En mettant en avant la pratique, nous cherchons à ramener au premier plan l'éthique. Contrairement à la tendance actuelle consistant à placer l'éthique en aval de la recherche produisant les savoirs théoriques et du développement des techniques (l'éthique interviendrait donc seulement au niveau des usages des outils et technologies), nous proposons de replacer l'éthique au cœur de l'ensemble de l'agir humain. Ce n'est cependant pas seulement (ou directement) une éthique *by design* qui est envisagée (nous y reviendrons plus loin). Ce que nous souhaitons encourager, c'est le développement des capacités éthiques qui s'exercent en amont du choix de faire ou non appel à une technique comme l'IA.

© 2025 Elsevier Masson SAS. Tous droits réservés. - Document téléchargé le 07/03/2025 par CH BELFORT MONTBELIARD - (97567). Il est interdit et illégal de diffuser ce document.

Les auteurs de cet article participent d'ailleurs depuis trois ans à un projet de recherche-action international expérimentant des manières de développer les capacités éthiques de citoyens. Ce projet s'intitule "Nouvel humanisme au temps des neurosciences et de l'intelligence artificielle" [4].

Machine qui devient intelligente, ou humain qui devient machine

Les progrès techniques réalisés depuis trois siècles dépendent tous d'une forme particulière d'intelligence : analytique, logicomathématique. En laissant ce type de rationalité dominer nos activités, nous assistons à la transformation réciproque de l'esprit humain et de son environnement de plus en plus artificialisé : la machinisation de notre environnement conduit en retour à la machinisation de notre manière de penser [5].

L'IA pourrait ainsi être considérée comme le fleuron de cette histoire moderne consistant à envisager les rapports à la nature sous l'angle d'une seule forme de rationalité humaine. *« Le discours dominant des multinationales de l'IA véhicule une vision extrêmement réductrice de l'humain et de la vie. Il tend à assimiler la machine à l'humain et l'humain à une machine comme une autre, selon les perspectives cognitiviste et physicaliste dominantes dans les technosciences. »* [6]

Comment faire, dès lors, pour non pas supprimer la technique, mais plutôt lui donner un rôle dans nos expériences relationnelles sans mécaniser nos esprits et nos mentalités ? *« L'assimilation de l'homme à une machine a des conséquences morales et sociales majeures. Si l'homme est une machine, fût-elle informatique, on peut en disposer comme on veut. Les machines étant des choses déterminées, il n'y a pas de droit des machines. On peut les conditionner, les rééduquer, les programmer, les faire travailler, les utiliser. L'homme-machine est un homme chosifié, privé de sa spécificité humaine. »* [7]

Dans le domaine de la santé comme dans tout autre, ce que nous appelons "progrès" n'est-il pas en train de suivre cette direction ? Si à chaque étape du processus de soins, une IA peut intervenir pour non seulement soulager le travail humain, mais aussi et surtout le rendre plus performant, n'assistons-nous pas à une chosification progressive de tout ce qui tenait auparavant lieu d'humain, avec ses vulnérabilités et dans sa complexité ?

Pareille évolution est aujourd'hui envisagée. Les auteurs Kerrie Holley et Siupo Becker la présentent comme le prochain objectif pour le système de la santé : permettre aux institutions de soins de tirer parti à grande échelle des avantages de l'IA, et plus uniquement en ce qui concerne certaines applications partielles. Ils encouragent la transformation de l'ensemble du système des soins de santé en mettant l'IA aux avant-postes, et ce afin d'améliorer les soins de santé pour tous : *« Le potentiel d'amélioration des soins aux patients existe lorsque l'utilisation de l'IA élimine ou rationalise les tâches non cliniques, donnant ainsi aux médecins plus de temps pour les soins directs aux patients et réduisant leur épuisement et leur surcharge. »* [2]

Ces mêmes auteurs déplorent que l'IA soit encore *« utilisée de manière discrète dans les soins de santé, en raison du scepticisme des patients et des prestataires et de leur résistance au changement »* [2]. Ils prônent cette médecine de surveillance que nous évoquions plus haut : *« Un nouveau paradigme informatique émergeant et les possibilités qui en découlent pour améliorer les soins de santé grâce au suivi en temps réel, à la surveillance continue et à la personnalisation extrême, tout en rendant les soins de santé plus accessibles, plus efficaces et plus utiles »* [2].

Face à ce risque de rationalisation extrême des pratiques médicales, nous souhaitons proposer une réflexion plus générale portant sur notre rapport à la technique. Ce que nous cherchons ainsi à mettre en évidence, c'est une posture consistant à dire à la fois "oui" et "non" aux apports technologiques. Nous pensons que pour préserver le caractère humain essentiel et irremplaçable de toute relation de soins, il est important de développer nos capacités éthiques. Ces dernières sont indispensables au niveau de la pratique délibérative inhérente à toute rencontre entre des acteurs humains.

La technicisation de l'humanité

Interroger le rôle réservé aux outils d'IA dans les pratiques médicales nécessite de nous pencher sur le rôle donné aux techniques en général.

Nous envisagerons ici la technique en suivant l'approche proposée par Bruno Latour qui critique toute idée d'une séparation nette et radicale entre les humains et les techniques qu'ils utilisent. Ce qui agit n'est jamais un agent humain isolé, mais plutôt un acteur-réseau :

© 2025 Elsevier Masson SAS. Tous droits réservés. - Document téléchargé le 07/03/2025 par CH BELFORT MONTBELIARD - (97567). Il est interdit et illégal de diffuser ce document.

un réseau d'actants humains et non humains (un schéma actanciel). Un agent humain (par exemple un médecin) n'agit donc qu'en réseau avec d'autres actants, humains (d'autres médecins, des patients, etc.) et non humains (divers dispositifs techniques, par exemple des systèmes d'IA). Une technique devient donc un actant à part entière en étant enrôlée dans un schéma actanciel. Son insertion dans le schéma n'est pas neutre, mais transforme le réseau. Elle peut par exemple modifier la situation par *« la création d'un nouvel objectif qui ne correspond au programme d'action d'aucun des agents »* [8].

Ce qu'il importe de souligner c'est que faire appel à une technique pour obtenir un résultat donné (résoudre un problème, répondre à une question, atteindre un but) entraîne une modification du réseau (de nouveaux effets surgissent). L'objectif atteint n'est plus celui qu'avait visé au départ un agent seul. Latour utilise le terme "translation" pour parler de cette transformation de résultat, et précise sa pensée comme suit : *« J'entends par translation le déplacement, la dérive, l'invention, la médiation, la création d'un lien qui n'existait pas auparavant et qui modifie dans une certaine mesure deux éléments ou agents. »* [8]

Concrètement, l'utilisateur ne sort pas indemne de sa décision de faire appel à une technique pour réaliser (mieux, plus vite, etc.) son action. Dorénavant ce ne sera plus "son" action, et il deviendra lui-même "autre".

Si nous abordons selon cette perspective le recours à l'IA dans une pratique médicale, nous voyons apparaître plusieurs points d'attention :

- la responsabilité de l'action doit être partagée entre les actants humains et non humains (techniques). L'humain n'est pas seul à générer l'action, dont une partie ne relève plus seulement de lui et de sa pure volonté. C'est bien le réseau d'actants, un collectif hybride sociotechnique, humain et non humain, qui agit ;
- une incertitude (imprévisibilité) surgit forcément quant aux effets de l'action menée par un tel collectif hybride : le nouveau résultat échappe à l'intention de l'agent humain agissant seul ;
- face à ce réel non maîtrisable, nous ne pouvons qu'insister avec vigueur sur ce qui se joue en amont de la décision de faire appel ou non à un partenaire technique. Qu'est-ce qui va être transformé dans la pratique de cet amont ? Qu'est-ce qui doit être préservé ? Le deuxième point ci-dessus nous rappelle que cette décision doit être posée sans savoir avec certitude où mènera cette alliance avec une nouvelle technique.

La principale leçon que nous tirons de cette approche est que l'éthique et la pratique sont toujours premières lorsque l'on envisage de faire appel à des techniques, notamment des techniques d'IA. Cela signifie qu'il est illusoire, et même dangereux, de se fier aveuglément à des prédictions, et à toute forme de futurologie calculatrice guidée par des théories déracinées de leur site de production (pratique) d'origine. Le théorique, qui lui-même provient en effet toujours d'une pratique originaire, ne peut prendre intégralement le pas sur la pratique : ces deux-là devraient au contraire toujours travailler main dans la main.

Il est illusoire, et même dangereux, de se fier aveuglément à des prédictions, et à toute forme de futurologie calculatrice

Il ne convient dès lors pas de dire que la technique n'est ni bonne ni mauvaise, mais bien plutôt qu'elle peut être bonne et mauvaise, étant donné qu'elle ne manque pas de transformer l'action et d'induire des effets inattendus. Ce sont les acteurs humains seuls qui peuvent s'assurer qu'elle sera la plus bénéfique possible pour le patient. La technique apportera, on l'espère, du bon, mais aussi des difficultés à ne pas occulter, et à prendre en charge. Dans le contexte de la pratique médicale, c'est-à-dire celui où les actions prennent place à l'intérieur de la relation faisant se rencontrer des soignants et des patients, le recours à la technique et l'engagement dans des collectifs hybrides méritent donc la plus grande attention.

L'informaticien italien Federico Cabitza défend un point de vue qui rejoint le nôtre [9]. Pour lui, la technologie n'existe en effet qu'en tant que reliée à l'utilisation pratique que nous en faisons. Il affirme ainsi que nous ne voyons en réalité pas la technique lorsque nous l'utilisons. Ce que nous voyons, ce sont ce que nous appelions plus haut des collectifs hybrides apparaissant sous la forme de projets, d'actions, de comportements, de dispositions, d'habitudes,

© 2025 Elsevier Masson SAS. Tous droits réservés. - Document téléchargé le 07/03/2025 par CH BELFORT MONTBELIARD - (97567). Il est interdit et illégal de diffuser ce document.

de compétences, bref de pratiques que les techniques rendent possibles en s'hybridant avec les agents humains. Cabitza propose ainsi de définir la technique comme un ensemble de règles et de ressources déployées dans la pratique.

Pour lui encore, l'IA peut être envisagée comme une technique rendant possible l'automatisation de tâches qui, normalement, nécessitent une certaine intelligence pour qu'elles puissent être accomplies par un être humain. En partant de cette définition de l'IA, l'auteur soutient qu'elle n'est pas véritablement intelligente parce que l'intelligence réside dans le fait de choisir une certaine tâche, et non dans le fait d'automatiser la tâche en question *via* des codes et des algorithmes. De manière à développer ses idées, l'informaticien reprend au futuriste allemand Gerd Leonhard le terme "androrithme" qui désigne toutes les caractéristiques spécifiques des êtres humains nous rendant irréductibles à des algorithmes : par exemple, notre désir d'anonymat et de confidentialité, le plaisir que nous procurent la sérendipité et l'imprévu, notre tendance à être anxieux, douteux, têtus, incohérents ou ambigus, notre droit à ne pas être trop méticuleux, notre besoin de silence, et ainsi de suite. Il soutient qu'il faut reconnaître la centralité des androrithmes dans l'existence humaine, ce qui signifie accepter la pluralité des êtres humains ainsi que l'incertitude de notre vie.

Nous nous rallions aussi à sa proposition d'adopter une approche éthique qu'il appelle *ethics in design* pour souligner qu'elle ne se focalise pas sur les algorithmes, mais plutôt sur les habitudes et les valeurs éthiques des humains qui créent, analysent, testent et valident les techniques numériques. L'éthique *in design* signifie donc que les praticiens (informaticiens, médecins, etc.) doivent être responsabilisés quant aux risques inhérents aux différentes étapes du recours à une technique, et en particulier à une IA. Cette éthique dans la pratique ne peut suivre une approche trop déontologique (principalisme), mais se rapproche au contraire d'une conception pragmatiste cultivant l'attention aux effets et prenant soin d'eux.

Conclusion

Au moment de clore cette réflexion, lisons encore quelques mots symptomatiques de notre époque, écrits par des auteurs déjà cités : *« La correspondance entre génotype et phénotype reste le plus vexant des mystères de la biologie ; mais les nouvelles méthodes d'apprentissage statistique offrent un contournement. Elles permettent d'automatiser l'analyse de jeux de données moléculaires extraordinaires. »* [1]

Que ce soit au niveau moléculaire ou à tout autre niveau, les données sont devenues le vaste terrain de jeu des IA et de leur toute-puissance statistique, permettant de débusquer toujours plus de corrélations et de régularités autrement inaccessibles aux humains, notamment aux médecins.

Pourtant, si la relation entre génotype et phénotype reste un mystère, n'est-ce pas pour la simple et bonne raison qu'en elle s'ouvre tout un espace de pratiques ? Et n'est-ce pas dans l'épaisseur de cet espace que l'activité relationnelle du prendre soin trouve tout son sens ? Pour demeurer dans leur art – terme rendant bien compte de la primauté de la pratique –, les médecins devraient selon nous rester les coauteurs, avec leurs patients, des choix qui s'opèrent lorsqu'est envisagé un recours à des techniques. Si une IA est introduite, les effets désirés et recherchés (réparer, guérir, soigner, soulager, etc.) devront faire l'objet de négociations entre les partenaires de ce collectif hybride.

Même si l'IA promet davantage de précision, d'efficacité, et de solutions personnalisées, répondant ainsi parfaitement à l'impératif biomédical d'éviter l'erreur, elle ne peut s'adresser directement au patient dont la singularité est au-delà de tout calcul. Ce que réclame le patient, c'est une rencontre qui soit capable d'accueillir sa vulnérabilité plutôt que de la réduire à une donnée statistique ou de la comparer à une norme étrangère.

Reconnaître que la pratique médicale est avant tout une affaire humaine, qui concerne la santé humaine (et la santé des vivants interdépendants), entraîne des précautions particulières au niveau des choix de faire appel ou non à des techniques. Rappelons-le, ces choix relèvent d'une dimension pratique et éthique d'amont. Ainsi, l'"IA first" prônée par certains auteurs à l'échelle du système de santé en sa globalité nous paraît déraisonnable (presque une contradiction dans les termes) : la technique ne peut être introduite qu'en aval de la pratique vécue dans les rencontres humaines. ■

Références

[1] Nordlinger B, Villani C. Introduction. In: Nordlinger B, Villani C (dir), editors. Santé et intelligence artificielle. Paris: CNRS Éditions; 2018.
[2] Holley KL, Becker S. AI-First healthcare. Sebastopol (États-Unis): O'Reilly Media; 2021.
[3] Zuboff S. Age of surveillance capitalism: the fight for a human future at the new frontier of power, public affairs. New York (États-Unis): PublicAffairs; 2019. Paris: Éditions Zulma; 2020 pour la traduction française.
[4] Nouvel humanisme au temps des neurosciences et de l'intelligence artificielle. Projet NHNAI. https://nhnai.org/fr/.
[5] Laurent N, Giorgi F, Guillermin M. Dire oui et non aux outils d'intelligence artificielle pour préserver la relation de soins. Ethica Clinica 2024;113.
[6] Ravet JC. Le devenir machine de l'être humain ? Relations 2020;(808):26–7.
[7] Juignet P. De l'homme-machine au cerveau-machine. Philosophie, science et société. 21 avril 2015. https://philosciences.com/homme-machine-cerveau-machine.
[8] Latour B. On technical Mediation. Common Knowledge 1994;3(2):29–64.
[9] Cabitza F. Deus in machina? L'uso umano delle nuove macchine, tra dipendenza e responsabilità. In: Floridi L, Cabitza F (dir), editors. Intelligenza artificiale. L'uso delle nuove macchine. Milan (Italie): Bompiani; 2021. p. 8–63.

Déclaration de liens d'intérêts
Les auteurs déclarent ne pas avoir de liens d'intérêts.

© 2025 Elsevier Masson SAS. Tous droits réservés. - Document téléchargé le 07/03/2025 par CH BELFORT MONTBELIARD - (97567). Il est interdit et illégal de diffuser ce document.