

**CONCOURS OUVERTS LES 10, 11, 12 et 13 JUIN 2025
POUR L'ADMISSION AU CYCLE DE FORMATION DES
ELEVES DIRECTEURS D'HÔPITAL**

4^{ème} EPREUVE D'ADMISSIBILITE

**(Durée 4 heures – Coefficient 4 si choisie comme majeure /
Coefficient 2 si choisie comme mineure)**

Vendredi 13 juin 2025

ECONOMIE DE LA SANTE

CONCOURS INTERNE et 3^{ème} CONCOURS

Sujet :

Vous êtes Directeur(rice) des affaires générales d'un Centre Hospitalier. Votre Directeur vous demande de lui transmettre une note opérationnelle présentant les avantages et les risques de l'intelligence artificielle en santé, et de dessiner les éventuels effets sur l'emploi des professionnels de santé.

Liste de documents :

Document 1 : INSERM - *Biais cognitifs dans le soin : comment l'IA générative pourrait aider à améliorer la prise en charge* - 20 février 2025. **(2 pages)**

Document 2 : OMS - *L'OMS publie le premier rapport mondial sur l'intelligence artificielle (IA) appliquée la santé et six principes directeurs relatifs à sa conception et à son utilisation* - 28 juin 2021 **(3 pages)**

Document 3 : BPI France - *IA dans le secteur de la santé : les 5 avantages clés* - 23 mai 2024 **(3 pages)**

Document 4 : Politics and Rights Review - *L'intelligence artificielle et son impact sur la santé, l'emploi et la croissance économique* - 8 janvier 2025 **(3 pages)**

Document 5 : Institut Montaigne - *IA et emploi en santé : quoi de neuf docteur ?* - Janvier 2019 **(2 pages)**

Document 6 : Le Monde - *Comment l'IA bouscule le milieu de la santé mentale* - 06 août 2024 **(2 pages)**

Biais cognitifs dans le soin : comment l'IA générative pourrait aider à améliorer la prise en charge

20 Fév 2025 | Par Inserm (Salle de presse) | Santé publique

Les biais cognitifs humains peuvent particulièrement impacter la prise de décision lorsque celle-ci doit être rapide, en particulier lorsqu'elle présente un enjeu vital, comme lors de la prise en charge médicale aux urgences par exemple. Une équipe de recherche de l'Inserm et de l'université de Bordeaux a mis à l'épreuve une méthode avancée d'intelligence artificielle générative^[1], entraînée avec les données des dossiers patients correspondants à 480 000 entrées aux urgences du CHU de Bordeaux. Ses résultats, présentés lors de la conférence *Machine Learning for Health* à Vancouver et publiés en parallèle dans la revue *Proceedings of Machine Learning Research*, montrent que l'IA testée est susceptible de reproduire et de mesurer les biais des soignants relatifs au genre des patients lors du triage. Ils constituent un cas d'usage de la façon dont les nouveaux algorithmes d'IA générative peuvent être mis à profit pour identifier et comprendre les biais cognitifs humains.

Dans les situations d'urgence de soin nécessitant une prise de décision rapide, les biais cognitifs humains, en particulier ceux dits « de jugement », peuvent avoir un impact critique sur la décision médicale et sur le pronostic du patient. Ces « raccourcis cognitifs » interviennent en effet lorsque les personnes doivent former une opinion ou prendre une décision à partir d'une information incomplète ou peu nuancée. La prise de décision peut ainsi être affectée de manière inconsciente par ces biais (liés par exemple au sexe/genre, à l'âge, à l'ethnie...), et conduire à sous-estimer ou à surestimer la sévérité de l'état d'une personne.

Alors, comment mieux identifier ces biais et diminuer leur impact ? Une réponse pourrait se trouver dans l'intelligence artificielle et en particulier dans les IA génératives dites « grands modèles de langage » (LLMs), c'est-à-dire capables d'imiter les prises de décision humaine grâce à leur maîtrise du langage humain (à l'image de ChatGPT par exemple). Ces modèles sont en effet capables d'appréhender efficacement le « texte libre »^[2] qui constitue une grande partie des données cliniques relevées par les soignants, en particulier aux urgences hospitalières.

Une équipe dirigée par Emmanuel Lagarde^[3], directeur de recherche Inserm, au sein du centre de recherche Bordeaux Population Health (Inserm/Université de Bordeaux), s'est ainsi intéressée au potentiel de ces LLMs pour détecter et quantifier les biais de genre en situation de prise de décision rapide. Le choix du contexte d'évaluation de cette méthode s'est porté sur le cas spécifique du triage^[4] des patients aux urgences médicales. La justesse de ce dernier est en effet critique : la sous-estimation d'une urgence qui reporterait la prise en charge peut entraîner la dégradation du pronostic d'un patient. A contrario, surestimer la gravité de l'état de la personne peut entraîner une surutilisation de ressources qui peut être particulièrement préjudiciable en cas d'affluence forte.

Les scientifiques ont utilisé une approche innovante, consistant à entraîner l'IA à trier les patients à partir des textes contenus dans leur dossier, reproduisant ainsi les éventuels biais cognitifs du personnel infirmier en charge de ce triage. Cet apprentissage du modèle a porté sur les dossiers de plus de 480 000 entrées au service des urgences du CHU de Bordeaux entre janvier 2013 et décembre 2021.

Une fois entraîné, le modèle était capable d'attribuer un score de triage (évaluant la sévérité de l'état du patient ou de la patiente) à partir de la lecture d'un dossier, comme le ferait le soignant. Le dossier était ensuite maquillé, de façon à modifier le genre de la personne dans les textes cliniques, et un nouveau score était attribué par le modèle. C'est la différence entre ces deux notations, produites à partir du dossier original d'une part et du dossier maquillé d'autre part, qui a ensuite permis l'estimation du biais cognitif.

Inserm salle de presse

Les résultats ont montré un biais significatif de l'IA au détriment des femmes : à dossiers cliniques identiques, la sévérité de leur état avait tendance à être sous-évaluée par rapport à celle des hommes (environ 5 % étaient classées « moins critiques » tandis que 1,81 % étaient classées comme « plus critiques »). A contrario, la sévérité de l'état des hommes avait tendance à être légèrement surévaluée (« plus critiques » pour 3,7 % contre 2,9 % « moins critiques »). Ce biais était d'autant plus marqué que le personnel infirmier était inexpérimenté.

« Ces travaux montrent comment les grands modèles de langage peuvent aider à détecter et à anticiper les biais cognitifs humains, ici dans l'exemple d'un objectif de prise en charge plus équitable et plus efficace aux urgences médicales », précise Emmanuel Lagarde. « La méthode utilisée montre que, dans ce contexte, les modèles LLMs sont capables d'identifier et de reproduire les biais qui guident la prise de décision humaine dans les données cliniques relevées par le personnel soignant », ajoute Ariel Guerra-Adames, doctorant et premier auteur de ces travaux[5].

Les prochains travaux de l'équipe vont se concentrer maintenant sur l'évaluation des biais liés à d'autres caractéristiques des patients (âge, groupe ethnique). À terme, le système devrait également être affiné avec l'introduction de variables non verbales (expressions faciales, ton de la voix) qui n'apparaissent pas nécessairement dans les données écrites et peuvent être pourtant critiques dans la prise de décision.

[1] L'intelligence artificielle générative est un système d'IA capable de créer du contenu, qu'il s'agisse de textes, d'images, de sons, de vidéos ou d'autres formes de données.

[2] En contexte médical, le texte libre désigne des informations consignées sous forme de texte non structuré, c'est-à-dire sans organisation rigide ou format prédéfini. Cela inclut les écrits rédigés directement par les professionnels de santé pour décrire des observations, diagnostics, traitements ou antécédents, souvent dans un langage naturel.

[3] En collaboration avec Cédric Gil-Jardiné du service des urgences du CHU de Bordeaux et Marta Avalos du centre Inria de l'université de Bordeaux

[4] Le triage aux urgences médicales consiste à classer les patients en fonction de la sévérité de leur état, afin d'optimiser l'ordre de prise en charge et ainsi de sauver un maximum de personnes. Il est réalisé par des personnels infirmiers dédiés qui, pour ce faire, collectent auprès de chaque malade différentes informations (raison de la visite, signes vitaux, historique médical...) et attribuent un score « d'urgence » selon une échelle validée.

[5] Ariel Guerra-Adames a reçu le prix de la meilleure communication dans le cadre de la présentation de ce travail pour la conférence *Machine Learning for Health* à Vancouver.



L'OMS publie le premier rapport mondial sur l'intelligence artificielle (IA) appliquée à la santé et six principes directeurs relatifs à sa conception et à son utilisation

Le recours croissant à l'IA dans le domaine de la santé présente des opportunités et des défis pour les gouvernements, les prestataires et les communautés

28 juin 2021

Communiqué de presse

Selon de nouvelles orientations publiées aujourd'hui par l'OMS, l'intelligence artificielle (IA) constitue un grand espoir pour améliorer la prestation des soins et la médecine dans le monde entier, mais à condition de placer l'éthique et les droits humains au cœur de sa conception, de son déploiement et de son utilisation.

Le rapport consacré à l'éthique et à la gouvernance de l'Intelligence artificielle dans le domaine de la santé (Intitulé en anglais : *Ethics and governance of artificial intelligence for health*) est le résultat de deux années de consultations menées par un groupe d'experts internationaux nommés par l'OMS.

« Comme toute nouvelle technologie, l'intelligence artificielle présente un immense potentiel pour améliorer la santé de millions de personnes dans le monde, mais comme pour toute technologie, il peut aussi en être fait mauvais usage et elle peut entraîner des effets préjudiciables, a déclaré le Dr Tedros Adhanom Ghebreyesus, Directeur général de l'OMS. Ce nouveau rapport, d'une importance capitale, constitue un guide précieux pour les pays qui souhaitent maximiser les avantages de l'IA, tout en minimisant les risques et en évitant les pièges. »

L'intelligence artificielle peut servir, comme c'est déjà le cas dans certains pays riches, à améliorer la rapidité et la précision du diagnostic et du dépistage des maladies, à faciliter les soins cliniques, à renforcer la recherche dans le domaine de la santé et la mise au point de médicaments, ainsi qu'à soutenir diverses interventions de santé publique, comme la surveillance des maladies, la riposte aux flambées épidémiques et la gestion des systèmes de santé.

L'IA pourrait également permettre aux patients de mieux maîtriser les soins de santé dont ils bénéficient et de mieux comprendre l'évolution de leurs besoins. Elle pourrait faciliter l'accès aux services de santé dans les pays disposant de peu de ressources et dans les communautés rurales, où les patients peinent souvent à accéder aux agents de santé ou au personnel médical.

Toutefois, ce nouveau rapport de l'OMS demande de ne pas surestimer les avantages de l'IA dans le domaine de la santé, en particulier lorsque cette technologie vient supplanter des investissements et des stratégies de base nécessaires pour parvenir à la couverture sanitaire universelle.

Il souligne également que les opportunités sont liées à des défis et à des risques, notamment une collecte et une utilisation contraires à l'éthique de données relatives à la santé, les biais introduits dans les algorithmes et les risques pour la sécurité des patients, la cybersécurité et l'environnement.



Par exemple, si les investissements privés et publics dans le développement et le déploiement de l'IA sont essentiels, l'utilisation non réglementée de cette technologie pourrait subordonner les droits et les intérêts des patients et des communautés aux puissants intérêts commerciaux des entreprises technologiques ou à ceux des gouvernements pour ce qui est de la surveillance et du contrôle social.

Le rapport souligne également que les systèmes entraînés principalement à partir de données recueillies auprès d'individus vivant dans des pays à revenu élevé risquent de ne pas bien fonctionner pour les populations des pays à revenu faible ou intermédiaire.

Les systèmes d'IA doivent donc être soigneusement conçus afin qu'ils tiennent compte de la diversité des contextes socio-économiques et de soins de santé. Ils doivent s'accompagner d'une formation aux compétences numériques, de la participation de la communauté et d'une sensibilisation, en particulier pour les millions d'agents de santé dont l'automatisation des rôles et des fonctions leur demandera d'acquérir des connaissances numériques ou de les perfectionner, et qui devront composer avec des machines susceptibles de remettre en question la prise de décision et l'autonomie des prestataires et des patients.

En fin de compte, guidés par les lois et obligations existantes en matière de droits humains, ainsi que par les nouvelles lois et politiques qui garantissent des principes éthiques, les gouvernements, les prestataires et les concepteurs doivent collaborer de sorte à tenir compte des problèmes d'éthique et de droits humains à chaque étape de la conception, du développement et du déploiement d'une technologie s'appuyant sur l'intelligence artificielle.

Six principes pour que, dans tous les pays, l'IA œuvre dans l'intérêt public

Afin d'atténuer les risques et de maximiser les opportunités intrinsèques à l'utilisation de l'IA dans le domaine de la santé, l'OMS propose que la réglementation et la gouvernance y afférentes reposent sur les principes suivants :

Protéger l'autonomie de l'être humain. Dans le contexte des soins de santé, les individus doivent rester maîtres des systèmes de soins de santé et des décisions médicales, la vie privée et la confidentialité doivent être protégées et les patients doivent donner un consentement éclairé valide au moyen de cadres juridiques appropriés en matière de protection des données.

Promouvoir le bien-être et la sécurité des personnes ainsi que l'intérêt public. Les concepteurs de technologies d'IA doivent respecter les obligations réglementaires relatives à la sécurité, à la précision et à l'efficacité pour des utilisations ou des indications bien définies. Il faut pouvoir disposer de mesures de contrôle de la qualité dans la pratique et d'amélioration de la qualité dans l'utilisation de l'IA.

Garantir la transparence, la clarté et l'intelligibilité. La transparence exige que des informations suffisantes soient publiées ou documentées avant la conception ou le déploiement d'une technologie d'IA. Ces informations doivent être facilement accessibles et permettre une consultation et un débat publics constructifs sur la conception de la technologie et sur l'utilisation qui doit ou non en être faite.



Encourager la responsabilité et l'obligation de rendre des comptes. Même si les technologies d'IA permettent d'accomplir des tâches spécifiques, il incombe aux parties prenantes de veiller à ce qu'elles soient utilisées dans des conditions appropriées et par des personnes dûment formées. Des mécanismes efficaces doivent être mis en place pour permettre aux individus et aux groupes lésés par des décisions fondées sur des algorithmes de contester ces décisions et d'obtenir réparation.

Garantir l'inclusion et l'équité. L'inclusion suppose que l'IA appliquée à la santé soit conçue de manière à encourager l'utilisation et l'accès équitables les plus larges possibles, indépendamment de l'âge, du sexe, du genre, des revenus, de la race, de l'origine ethnique, de l'orientation sexuelle, des capacités ou d'autres caractéristiques protégées par les codes relatifs aux droits humains.

Promouvoir une IA réactive et durable. Les concepteurs, les développeurs et les utilisateurs devraient évaluer de manière continue et transparente les applications de l'IA en situation réelle afin de s'assurer que cette technologie répond de manière adéquate et appropriée aux attentes et aux besoins. Les systèmes d'IA devraient également être conçus de sorte à réduire au minimum leurs conséquences environnementales et à accroître leur efficacité énergétique. Les gouvernements et les entreprises devraient anticiper les bouleversements qui seront occasionnés au niveau du travail, notamment la formation des agents de santé qui devront se familiariser avec l'utilisation des systèmes d'IA, et les pertes d'emploi que le recours à des systèmes automatisés est susceptible d'engendrer.

Ces principes guideront les travaux futurs de l'OMS en vue de garantir que le plein potentiel de l'IA en matière de soins de santé et de santé publique sera mis au service du bien de tous.

IA dans le secteur de la santé : les 5 avantages clés

De plus en plus d'entreprises de la Health Tech ont recours à l'intelligence artificielle. De la détection précoce des troubles au suivi des patients, l'IA a en effet le potentiel d'améliorer substantiellement le quotidien des patients et des professionnels de santé. Big média revient sur quelques-uns des nombreux bénéfices de l'IA dans ce domaine.

23 mai 2024

On entend parler d'elle dans tous les secteurs d'activité : transports, commerce, finance, environnement, défense, industrie... l'intelligence artificielle s'est imposée en quelques années comme un des tremplins technologiques majeurs des sociétés de demain. Si l'IA comporte son lot de risques, ce que vise d'ailleurs à encadrer le nouveau règlement européen sur l'intelligence artificielle (IA Act) adopté par le Conseil européen le 21 mai dernier, cette technologie profite indéniablement à de nombreux domaines – et la santé au premier chef. L'Organisation de coopération et de développement économiques (OCDE) relève en effet que le marché mondial de l'IA en santé pourrait être multiplié par 16 d'ici 2030 et passer ainsi de 11 milliards de dollars en 2021 à... 188 milliards de dollars¹. La France, de son côté, n'est pas en reste et semble particulièrement à la pointe de l'utilisation de l'IA dans la santé : d'après une étude de la medtech PulseLife², plus d'un soignant sur deux (53 %) aurait recours à l'IA dans sa pratique. De plus, le pays compte plus de 2 600 PME innovantes dans la santé, dont pas moins de 450 sociétés de numérique en santé et IA³ (le double d'il y a quatre ans).

Mais concrètement, qu'est-ce que l'IA présente comme avantages pour le secteur ?

L'IA comme aide à la détection, au diagnostic médical et au traitement d'une maladie

L'intelligence artificielle doit permettre d'affiner les diagnostics des médecins mais aussi, en amont, de détecter une pathologie, une maladie rare ou invisible et ainsi d'établir le traitement le plus approprié. Selon un sondage IFOP réalisé pour Sanofi, 81 % des médecins pensent que l'IA pourrait contribuer à l'amélioration du diagnostic des maladies rares. En entraînant des algorithmes de deep learning ou encore d'apprentissage automatique sur les grandes quantités de données produites par les systèmes de santé, des dossiers des patients jusqu'aux actes

d'imagerie médicale (plus de 96 millions générés en 2022 dans le secteur libéral⁴), les diagnostics pourraient être plus rapides, précis et personnalisés. Cela offrirait, de facto, plus de temps aux professionnels de santé pour le suivi de leurs patients. Les champs d'application sont d'ores et déjà nombreux : oncologie, ophtalmologie, cardiologie (à l'instar de la start-up santé Cardiologs) ou dermatologie, avec par exemple les startups françaises Therapixel et Damae Medical qui rendent respectivement possible, par l'implémentation de l'IA à leurs solutions, la détection précoce du cancer du sein et des cancers cutanés. L'IA présente également un intérêt certain pour la médecine génomique, qui manie d'importantes masses de données.

Protéger les systèmes de santé contre les menaces cyber

Les données de santé sont des données sensibles qu'il convient de protéger coûte que coûte. Or on ne compte plus les cyberattaques, notamment sur les hôpitaux. En avril dernier, l'hôpital de Cannes a par exemple subi une cyberattaque provoquant la fuite de 61 giga-octets de données. Sur les 581 incidents informatiques dans les hôpitaux et établissements médico-sociaux recensés en 2023 par l'Agence du Numérique en Santé, près de la moitié sont des cyberattaques. Le préjudice financier de ce type de malveillances pourrait atteindre, dans le monde, la somme colossale de 10 500 milliards de dollars⁵ d'ici 2025. Or il se pourrait que l'IA, à condition d'être utilisée à bon escient, puisse prévenir et contrer ce genre d'opérations. Détecter et bloquer les attaques en temps réel, anticiper un comportement suspect, renforcer la sécurité des données, sécuriser le réseau, voire prédire de futures failles... Si elle est parfois à la source du problème en matière de cyberattaque et suscite des inquiétudes à ne pas ignorer, l'IA semble également être, entre de bonnes mains, un outil efficace pour le renforcement de la sécurité des systèmes de santé.

Utilisation de l'IA par les professionnels de santé dans leurs interventions

On a vu que l'intelligence artificielle pouvait aider concrètement les professionnels dans leurs pratiques, par exemple dans la détection anticipée de pathologies ou d'anomalies. Mais l'IA intervient également dans l'acte de soin en lui-même : en chirurgie, notamment, les avancées sont particulièrement remarquables. L'IA peut être à tous les stades de l'opération, de la planification à la modélisation jusqu'à la réalisation. Certains hôpitaux de l'AP-HP ont par exemple fait entrer au bloc opératoire Connected Surgery, une plateforme d'assistance chirurgicale basée sur l'IA – laquelle peut également servir à améliorer la sécurité des patients durant une opération, à l'instar du CHU de Strasbourg, qui a intégré à son bloc opératoire un dispositif innovant permettant de prévenir les risques d'hypotension artérielle (qui ne peut pas être anticipée avec les outils traditionnels de surveillance de la pression artérielle) lors des interventions chirurgicales. L'intelligence artificielle n'est pas la seule technologie novatrice à venir en aide aux praticiens (casque de réalité mixte, robots-chirurgiens, impression 3D, etc.).

De nombreux hôpitaux ont déjà adopté ce type de dispositifs, qu'il convient évidemment d'encadrer d'un point de vue éthique et déontologique.

L'intelligence artificielle générative pour pousser la découverte de nouveaux médicaments

La découverte de médicaments est un processus coûteux et (très) chronophage. Ainsi la mise en vente d'un nouveau médicament peut-elle prendre des dizaines d'années (R&D, phase clinique, etc.) et coûter plusieurs milliards de dollars. Ces dernières années ont vu fleurir de nombreuses startups et « techbio » innovantes qui, associées à l'industrie pharmaceutique, misent sur l'intelligence artificielle générative pour accélérer drastiquement la recherche et le design de nouvelles molécules. C'est par exemple le cas des startups françaises Bioptimus, fondée par deux anciens de Google et Owkin, qui a récemment réalisé une levée de fonds de 35 millions de dollars, Iktos ou encore Agemia, dont le nouveau partenariat avec Sanofi s'élève à 140 millions de dollars. D'autres, comme la britannique Exscientia, se focalisent plus spécifiquement sur le design de médicaments, avec une solution d'IA générant des propositions de composés. Réduction des coûts, des délais, des échecs cliniques, optimisation des performances des molécules ou identification de nouvelles cibles thérapeutiques, l'IA a définitivement un rôle à jouer dans la découverte des médicaments de demain.

Utiliser l'IA pour veiller sur la santé mentale des patients

L'intelligence artificielle a également sa place dans le suivi des personnes souffrant de troubles mentaux. En 2021, l'Organisation mondiale de la santé (OMS) estimait que plus de 150 millions de personnes souffraient d'un problème de santé mentale en Europe, une situation qui ne s'est certainement pas arrangée avec le Covid-19. Or l'IA semble pouvoir appuyer – et en aucun cas remplacer – le travail des professionnels de santé dans la détection et le suivi de troubles mentaux. Cela peut passer par le ciblage de certaines catégories de populations particulièrement sujettes à la dépression, l'anxiété ou des troubles neuropsychiatriques : une deeptech comme Emobot a par exemple mis au point un petit robot équipé d'une IA capable d'anticiper les troubles de l'humeur chez les personnes âgées vivant en Ehpad, quand d'autres, telles que Présage ou Oso-Ai, permettent d'anticiper les situations de détresse ou les risques d'hospitalisations. Des recherches scientifiques s'aidant de l'IA ont également lieu en ce qui concerne la détection de troubles anxieux à l'adolescence, à l'instar de cette étude du laboratoire Trajectoires développementales et psychiatrie (Inserm/ENS Paris-Saclay) et du Centre Borelli (ENS/CNRS/Université Paris-Saclay). En analysant conjointement les données de la génétique, de l'épigénétique et de l'imagerie médicale, les outils d'intelligence artificielle pourraient globalement permettre aux chercheurs de constituer une connaissance plus précise des pathologies mentales et de détecter d'éventuels signes avant-coureurs.

Politics and Rights Review

L'intelligence artificielle et son impact sur la santé, l'emploi et la croissance économique
L'intelligence artificielle (IA) redéfinit les industries à l'échelle mondiale, offrant des avancées sans précédent tout en introduisant des défis que les sociétés, les organisations et les gouvernements doivent gérer avec prudence.

janvier 8, 2025

L'intelligence artificielle (IA) redéfinit les industries à l'échelle mondiale, offrant des avancées sans précédent tout en introduisant des défis que les sociétés, les organisations et les gouvernements doivent gérer avec prudence. Les capacités de l'IA sont particulièrement évidentes dans des domaines comme la santé, l'emploi, le comportement organisationnel et les systèmes économiques, où elle automatise des processus complexes, améliore la productivité et favorise des solutions innovantes.

Cependant, ces transformations ne sont pas sans poser des préoccupations éthiques, sociales et économiques importantes, notamment la suppression d'emplois, les risques pour la vie privée et une évolution des compétences requises sur le marché du travail. L'analyse suivante explore les impacts de l'IA sur des secteurs clés, en s'appuyant sur des recherches récentes pour comprendre les implications à long terme de l'IA pour les individus et la société.

Dans son essence, l'expansion de l'IA soulève des questions cruciales sur l'équilibre entre le progrès et les valeurs centrées sur l'humain. Par exemple, bien que l'IA dans le domaine de la santé puisse révolutionner les diagnostics et les plans de traitement, sa dépendance aux données personnelles soulève des problèmes de confidentialité des patients et de sécurité des données.

De même, le rôle de l'IA sur les lieux de travail introduit des gains d'efficacité et rationalise la prise de décision, mais pourrait également redéfinir l'autonomie des travailleurs et les structures d'emploi. Pour comprendre et relever ces défis, cette analyse examine la double nature de l'impact de l'IA dans quatre domaines critiques : la santé, l'emploi, la dynamique organisationnelle et la croissance économique.

IA dans le domaine de la santé : Avantages et préoccupations éthiques

Les applications de l'IA dans le domaine de la santé révolutionnent le secteur médical en soutenant le diagnostic, la planification des traitements et l'efficacité opérationnelle.

Le système CyberKnife est un dispositif robotique avancé alimenté par l'IA, conçu pour la radiothérapie de précision. Photo de Fredrik Frantzen (CC BY-ND).

Les modèles d'apprentissage automatique, par exemple, aident à diagnostiquer des maladies telles que le cancer et les troubles neurologiques avec une précision impressionnante, soutenant une prise de décision clinique plus rapide et efficace. Un modèle d'IA a notamment atteint un taux de précision de 93,8 % pour détecter les marqueurs associés au cancer colorectal, démontrant le potentiel de l'IA pour améliorer le dépistage précoce et les résultats des patients (Rahman et al., 2024).

Au-delà du diagnostic, l'IA contribue de manière significative à la logistique sanitaire, réduisant les temps d'attente des patients, optimisant les flux de travail et permettant aux professionnels de se concentrer davantage sur les soins aux patients que sur les tâches administratives. Par exemple, les algorithmes qui traitent les données d'admission des patients et priorisent les cas peuvent éviter les goulots d'étranglement dans les services d'urgence, améliorant ainsi la qualité du service et sauvant des vies (Morandini et al., 2023).

Cependant, l'IA dans le domaine de la santé n'est pas sans soulever de sérieuses préoccupations éthiques et opérationnelles :

- **Risques pour la confidentialité des données :** La dépendance à de grands ensembles de données, souvent sensibles, soulève des inquiétudes concernant la sécurité des données. Les violations impliquant des dossiers médicaux pourraient exposer les patients à des violations de la vie privée et,

dans certains cas, à un vol d'identité, nécessitant des protocoles rigoureux de protection des données (Shen & Zhang, 2024).

- **Responsabilité dans les décisions cliniques :** Les diagnostics assistés par l'IA introduisent des questions de responsabilité complexes, en particulier lorsque des erreurs ou des diagnostics incorrects se produisent. Par exemple, une erreur dans le diagnostic des arythmies cardiaques en raison des limites d'un algorithme pourrait avoir des conséquences graves, compliquant la responsabilité des prestataires de soins de santé et des développeurs d'IA (Hankins et al., 2023).

IA et l'emploi : Déplacement des emplois et opportunités émergentes

L'intégration rapide de l'IA dans diverses industries redéfinit les marchés du travail, en automatisant les tâches routinières et en générant une demande pour des compétences spécialisées.

Bien que l'automatisation promette une productivité accrue, elle introduit également le risque d'un **déplacement significatif des emplois**, en particulier dans les rôles peu qualifiés. De nombreuses organisations ont commencé à automatiser les tâches répétitives, ce qui, bien qu'améliorant l'efficacité, impacte de manière disproportionnée les employés occupant des fonctions administratives et opérationnelles (Hui et al., 2024).

À l'inverse, l'émergence de rôles centrés sur l'IA reflète la **demande croissante de travail hautement qualifié** dans des domaines tels que la science des données, l'apprentissage automatique et le développement de l'IA.

Cette demande d'expertise spécialisée indique un **changement sur le marché du travail vers des rôles nécessitant des compétences techniques et une grande adaptabilité**. Il convient de noter que les initiatives de montée en compétences et de reconversion sont essentielles, car elles dotent la main-d'œuvre des compétences nécessaires pour prospérer dans des environnements de plus en plus pilotés par l'IA (Deranty & Corbin, 2022).

Pour explorer plus en détail l'avenir de l'emploi dans un monde enrichi par l'IA, prenez en compte les points suivants :

- **Automatisation des tâches routinières :** L'adoption de l'IA pour les tâches répétitives dans des secteurs comme la fabrication et le service client représente un risque pour les personnes traditionnellement employées dans ces domaines (Starace et al., 2024).
- **Demande de rôles hautement qualifiés :** À mesure que l'IA s'intègre dans diverses fonctions, des postes nécessitant des compétences avancées en analyse et une expertise technique émergent, signalant une évolution vers des emplois spécialisés qui pourraient aggraver les écarts de compétences et les disparités salariales (Attard-Frost et al., 2024).

L'avenir du travail dépendra de plus en plus du développement dynamique de la main-d'œuvre, avec un accent sur l'apprentissage continu et des compétences adaptables pour minimiser les perturbations socio-économiques causées par l'adoption de l'IA.

Transformation organisationnelle par l'IA : Efficacité et mécanismes de contrôle

L'intégration de l'IA dans les pratiques organisationnelles transforme les flux de travail, la prise de décision et les systèmes de gestion. Les organisations qui exploitent l'IA signalent une augmentation de la productivité, avec des systèmes automatisant des tâches allant de la maintenance prédictive dans les environnements industriels au contrôle qualité dans la fabrication. En s'appuyant sur des décisions basées sur les données, les organisations peuvent améliorer la précision et la cohérence de leurs opérations (Yakhshiboyev & Ermetov, 2024).

Cependant, le rôle de l'IA dans la gestion algorithmique introduit de nouveaux défis concernant la confidentialité et l'autonomie des employés. Par exemple, les systèmes qui surveillent la productivité des employés via une analyse continue des données peuvent créer un environnement de surveillance constante, ce qui pourrait réduire la satisfaction au travail et augmenter le turnover. Les systèmes de gestion algorithmique, bien qu'efficaces, manquent souvent de transparence, soulevant des inquiétudes sur l'équité dans l'allocation des tâches et l'évaluation des performances (Zarifhonarvar, 2024).

Pour équilibrer les aspects positifs et négatifs de l'IA dans les environnements organisationnels, considérez :

- **Gains de productivité** : Les systèmes automatisés permettent aux organisations d'optimiser les ressources et d'améliorer la productivité, en particulier dans la fabrication et la logistique (Morandini et al., 2023).
- **Défis de la gestion algorithmique** : La supervision pilotée par l'IA peut éroder l'autonomie des travailleurs et réduire la transparence, affectant le moral des employés et augmentant les tensions sur le lieu de travail.

Les organisations sont encouragées à intégrer l'IA avec transparence, en veillant à ce que les systèmes d'IA complètent, plutôt que remplacent, l'expertise humaine, favorisant un environnement où l'IA contribue à améliorer les capacités humaines.

Implications économiques : Croissance, innovation et recommandations politiques

En tant que moteur de croissance économique, le potentiel de l'IA pour améliorer la productivité et favoriser l'innovation est considérable. Les prévisions économiques estiment que l'IA pourrait contribuer à hauteur de 13 000 milliards de dollars à la production économique mondiale d'ici 2030, principalement en automatisant des tâches complexes, en optimisant les processus de production et en rationalisant la prestation des services (Rahman et al., 2024).

Cependant, maximiser les avantages de l'IA tout en atténuant ses risques socio-économiques nécessite des cadres réglementaires équilibrés. Par exemple, des politiques conçues pour protéger les travailleurs occupant des postes vulnérables à l'automatisation sont essentielles pour une croissance économique inclusive. De même, des filets de sécurité sociale robustes peuvent soutenir les individus dans les phases de transition, tels que les assurances chômage ou les programmes de reconversion, aidant à combler l'écart alors que les économies s'adaptent aux changements induits par l'IA.

Les recommandations politiques pour l'intégration économique de l'IA incluent :

- **Favoriser une croissance inclusive** : Les politiques encourageant des pratiques de travail équitables et une inclusion économique peuvent aider à réduire les disparités, notamment à mesure que les rôles hautement qualifiés deviennent plus dominants.
- **Régulation des prises de décision par l'IA** : Assurer la transparence dans les décisions pilotées par l'IA en matière de recrutement et de gestion aidera à atténuer les risques de biais, créant un environnement de travail plus juste et équitable.

En résumé, établir des politiques qui soutiennent l'innovation tout en assurant des pratiques de travail équitables peut aider à aligner les progrès économiques induits par l'IA avec des objectifs sociaux plus larges. À mesure que l'influence de l'IA sur les systèmes économiques s'accroît, les gouvernements et les organisations doivent prendre en compte les dimensions sociales et éthiques des avancées technologiques.

Conclusions et orientations futures

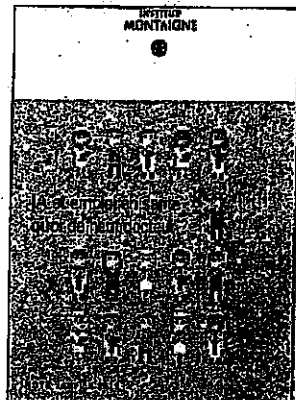
L'intégration de l'IA dans les domaines de la santé, de l'emploi, des organisations et des économies démontre son potentiel transformateur, bien que ces avancées nécessitent une gestion responsable.

Les cadres éthiques sont essentiels dans le domaine de la santé pour traiter les questions de confidentialité et de responsabilité. Dans l'emploi, les politiques qui soutiennent l'apprentissage continu et l'adaptation de la main-d'œuvre sont cruciales pour minimiser les risques de déplacement. Pour les organisations, la transparence dans la gestion algorithmique peut favoriser des environnements de travail plus équitables et inclusifs. Enfin, des politiques économiques équilibrées sont nécessaires pour garantir que les avantages de l'IA soient accessibles à tous.

Les recherches futures devraient se concentrer sur la compréhension des implications à long terme de l'IA sur les marchés du travail, en évaluant l'efficacité des cadres réglementaires et en explorant la relation évolutive entre les systèmes d'IA et le travail humain.

INSTITUT
MONTAIGNE

Résumé



En mars 2018, le Président de la République soulignait le caractère stratégique de l'Intelligence Artificielle (IA) en santé lors de la remise du rapport du député Cédric Villani « Donner un sens à l'intelligence artificielle ». Il annonçait que la santé serait

un secteur prioritaire pour le développement de l'IA. Ce thème était à nouveau repris lors de la présentation du Plan de transformation du système de santé « Ma Santé 2022 », qui comprend un important volet consacré au numérique.

Les applications de l'IA en santé sont multiples et auront des effets sur l'ensemble du parcours de soins : analyse des symptômes, pose d'un diagnostic, décision thérapeutique. Pour permettre le développement de solutions d'IA en santé performantes, l'accès aux données est un élément clé. Face à ces enjeux, les annonces récentes du rapport de la mission, destinée à préfigurer un *Health Data Hub* national vont dans le bon sens.

Si ces initiatives publiques visent à ouvrir la voie à l'IA en santé et à structurer un pilotage effectif pour les données de santé, elles n'intègrent pas, à ce stade, de stratégie d'anticipation et d'accompagnement des impacts de l'IA sur les métiers du secteur de la santé. L'élaboration d'une telle stratégie doit constituer une priorité afin d'anticiper et rendre plus efficace la transformation digitale du système de santé français.

Cette note a pour objectif de proposer une stratégie d'accompagnement des métiers du secteur de la santé et des fonctions support aux soins. Elle vise à contribuer à une prise de conscience quant à la nécessité de créer, dans les meilleurs délais, une méthodologie d'évaluation des effets du déploiement de l'IA et de la robotisation dans le secteur de la santé. Une telle méthodologie a pour objectif de rendre effective la transformation des métiers du soin.

Quels sont les effets de l'IA en santé ?

Dans les établissements de soins, les solutions d'IA les plus en pointe sont majoritairement présentes dans les centres les plus innovants et permettent de faire progresser la médecine et la recherche en santé. Ces technologies de pointe ne se substituent pas nécessairement à l'intervention humaine, mais augmentent les capacités du médecin (analyse, diagnostic, prescription). D'autres établissements se dotent de solutions plus généralistes, qui répondent à des besoins de santé courants et s'appuient sur la digitalisation de la prise en charge des patients. De tels dispositifs peuvent être déployés dans les territoires marqués par une pénurie d'offre médicale, contribuant à la recomposition de l'offre de soins en France.

Concernant les ressources humaines, les conditions d'exercice des spécialités médicales et des fonctions support aux soins vont durablement changer, même si les impacts diffèrent selon les métiers. Un impact possiblement fort est à attendre pour les disciplines médicales dont la matière première est déjà du code numérique comme la radiologie ou l'ophtalmologie. En conséquence, ces spécialités sont déjà impactées par l'IA. Leurs activités se déplaceront vers des actes à plus haute valeur ajoutée, tandis que les tâches les plus rébarbatives seront effectuées par la machine.

Si l'IA ne menace pas directement les médecins à court terme, ses effets seront plus lourds sur les fonctions support aux soins. Ainsi, plusieurs solutions d'IA en matière de gestion des tâches administratives ou logistiques, existent déjà dans d'autres secteurs.

Nos propositions

Proposition n°1 : établir une méthodologie de chiffrage des impacts de l'IA sur les ressources humaines en santé

La proposition d'établir une méthodologie d'évaluation des impacts de l'IA sur l'emploi permet d'appréhender les effets de création et de transformation des métiers du secteur de la santé. Les pouvoirs publics doivent créer une plateforme de dialogue entre les différents acteurs impliqués. La méthodologie de chiffrage proposée dans cette note s'articule autour de six étapes clés.

La première étape consiste à recenser les effectifs par catégories professionnelles du secteur de la santé, comme par exemple la gestion administrative. Ensuite, il s'agit de répertorier les métiers appartenant aux catégories professionnelles à l'instar de celui de secrétaire médical(e). Sont ensuite identifiées les activités ou tâches correspondant à chaque métier, ce qui permet, lors de la quatrième étape, de déterminer le taux de substitution de chaque activité identifiée. L'étape suivante vise à mesurer le taux de substitution pour chaque métier. Enfin, la dernière étape permet de réaliser des scénarios d'impact.

Il apparaît essentiel que les pouvoirs publics puissent rapidement se saisir de l'enjeu de construire une méthodologie de suivi en continu des impacts de l'IA sur l'emploi en santé. La construction d'un cadre de dialogue associant les représentants des patients, les professionnels de santé, les partenaires sociaux et les acteurs économiques constitue un axe prioritaire.

Proposition n°2 : déployer une RSE digitale au sein des établissements de santé

L'investissement dans l'anticipation du déploiement de l'IA et de la robotisation doit pouvoir être reconnu comme une nouvelle forme d'exercice de la Responsabilité Sociétale des Entreprises (RSE). Il s'agit d'appliquer la notion de RSE à l'engagement de mesures d'accompagnement des effets sociétaux induits par la transformation numérique et donc par l'IA. Le développement des technologies d'IA dans un secteur d'activité donné entraîne nécessairement des conséquences sur la société. Il convient donc de les encadrer de façon éthique et responsable.

La RSE digitale peut être déclinée autour de différentes mesures, telles qu'inciter les établissements de soins à étudier la transformation des métiers, les encourager à développer des méthodologies d'évaluation des compétences et des nouveaux besoins en formation des salariés. Ces derniers doivent être sensibilisés et formés à l'utilisation des innovations en santé. En ce qui concerne les patients, il s'agit de s'assurer du respect du principe de garantie humaine, soit la garantie d'une supervision humaine de toute utilisation du numérique et de l'IA en santé, pour toute décision médicale (pose de diagnostic, choix du traitement, etc.). Ces mesures concourent à la mise en place d'une IA responsable, en valorisant l'innovation à tous les niveaux.

Comment l'IA bouscule le milieu de la santé mentale : « Plutôt que de payer une nouvelle séance chez le psy, j'allais sur ChatGPT »

Publié le 06 août 2024

Décryptage - Selon OpenAI, son chatbot peut désormais tenir des discussions fluides à l'oral et « lire » les émotions de ses utilisateurs. Une innovation qui intervient alors que l'usage de l'intelligence artificielle générative à des fins thérapeutiques se répand, soulevant des interrogations chez les professionnels de santé.

« La meilleure expérience que j'ai eue avec la thérapie. » « J'ai pleuré plusieurs fois et j'ai eu beaucoup de révélations. Juste une journée à lui parler m'a fait me sentir mieux. » « Je lui ai confié quelque chose que je n'avais encore jamais dit à personne. » Ces témoignages diaphaniques ne sont pas extraits de la page Google d'un excellent thérapeute. Ils ne font même pas référence au travail d'un humain mais à deux robots conversationnels, des chatbots, créés sur la plate-forme américaine Character.ai.

Sur le site communautaire Reddit, plusieurs utilisateurs chantent ainsi les louanges de Therapist et de Psychologist, des robots « serviables », « compatissants » et qui donnent « l'impression de parler à un vrai thérapeute, mais un bon et raisonnable ». Ils sont également disponibles 24 heures sur 24 et gratuits, ce qui peut être un avantage lorsque l'on est « fâché », comme l'explique une internaute sur le forum.

La plate-forme Character.ai, dont les utilisateurs ont majoritairement entre 16 et 30 ans, compte environ 475 bots faisant office de thérapeutes, selon un décompte de la BBC en janvier. Le succès de certains de ces « psychologues » virtuels se compte en plusieurs centaines de milliers de visites.

Parmi eux, Psychologist est le plus célèbre. Il cumule à ce jour plus de 154 millions de conversations depuis sa création, il y a un peu plus d'un an. « Je n'ai jamais eu l'intention qu'il devienne populaire, ni que d'autres personnes l'utilisent comme un outil, a expliqué son créateur, un étudiant en psychologie néo-zélandais qui affirme avoir développé l'outil pour ses besoins personnels, car il n'avait pas les moyens de payer ses séances de thérapie. J'ai commencé ensuite à recevoir beaucoup de messages d'utilisateurs me disant qu'ils avaient été positivement affectés et qu'ils l'utilisaient comme source de réconfort. »

« Effet Eliza »

C'est le cas de Charlotte, 30 ans. Elle, c'est un autre outil d'intelligence artificielle (IA) générative, ChatGPT, qu'elle a utilisé en complément de ses séances de psychothérapie, l'été dernier. « Il m'arrivait de sortir de ma consultation et d'avoir encore des questions à poser, donc plutôt que de payer une nouvelle séance, ce qui coûte cher, je les posais directement à ChatGPT », raconte la trentenaire résidant en région parisienne. Avant de préciser : « Je ne prenais jamais pour argent comptant ce qu'il me disait, je voyais plus ça comme des pistes de réflexion. Un peu comme quand un ami te donne des conseils, mais que tu ne vas pas forcément appliquer tout ce qu'il te dit. » Elle se remémore avoir été « assez surprise » par la justesse de la réponse à sa première question, ce qui l'a rapidement convaincue d'y passer plus de temps : « C'est tout un univers qui s'ouvrait à moi, je pouvais facilement y passer une heure. »

Sauf que les réponses apportées par les chatbots « psychologues » sont parfois si convaincantes qu'elles peuvent avoir des effets tragiques. En mars 2023, un chercheur belge d'une trentaine d'années, victime d'écoanxiété, s'est donné la mort après six semaines de conversations intensives avec un chatbot nommé Eliza, de la société américaine Chai Research. « C'était comme une drogue dans laquelle il se réfugiait, matin et soir, et dont il ne pouvait plus se passer », a raconté sa femme au quotidien La Libre.

L'entreprise a assuré par la suite travailler à l'amélioration de la sécurité de son application : aujourd'hui, quand un utilisateur évoque par exemple le sujet du suicide avec le chatbot, celui-ci refuse de développer une réponse. Sans toutefois renvoyer automatiquement vers un professionnel de la santé mentale, contrairement à ChatGPT.

Cette dissonance cognitive subtile, consistant à assimiler de manière inconsciente le comportement d'une machine à celui d'un être humain, a un nom : « l'effet Eliza », justement, du nom du programme informatique considéré comme l'un des premiers robots conversationnels, écrit dans les années 1960 par Joseph Weizenbaum, informaticien au MIT (institut de technologie du Massachusetts). En simulant une psychothérapie de l'approche rogérianne, consistant à reformuler les réponses du patient sous forme de questions, ce programme a mis en évidence la tendance de ses utilisateurs à prêter une implication émotionnelle de la machine dans le processus de « thérapie ».

Exercice de respiration et analyse des émotions

Loin de le voir comme un défaut, certaines entreprises cherchent plutôt à capitaliser sur cet effet Eliza. En témoigne le dernier gadget dopé à l'IA présenté fin juillet, baptisé Friend (« ami »), qui prend la forme d'un médaillon ayant vocation à entretenir la conversation avec son porteur. Ou l'intégration par OpenAI d'un nouveau mode vocal dans la dernière version de son chatbot ChatGPT, GPT-4o. Celui-ci est désormais capable de tenir des conversations fluides à l'oral et de déchiffrer les émotions de ses utilisateurs, en se fondant sur l'expression de leur visage ou l'intonation de leur voix, assure l'entreprise. Le 13 mai, lors de la démonstration de l'outil – dont l'accès est pour l'heure limité à un petit nombre d'utilisateurs –, un salarié d'OpenAI allait même jusqu'à lui demander un exercice de respiration pour qu'il « se sente un peu moins nerveux ».

Comment les professionnels de la santé mentale vivent-ils cette incursion de l'IA sur leur terrain ? « Heureusement, notre travail en tant que psychologues va beaucoup plus loin que ça, et je ne pense pas que cela suffira à nous remplacer », répond Olivier Duris, psychologue clinicien et spécialiste des usages numériques. Pour lui, le problème se situe ailleurs : « Si les gens se tournent vers la machine, c'est parce qu'ils ont de moins en moins la possibilité de le faire avec des professionnels de santé. Ainsi, on se retrouve avec des entreprises comme OpenAI qui vous donnent l'illusion d'avoir trouvé la solution à vos problèmes, gratuite et rapide, alors qu'il ne s'agit que d'une simulation. »

Les promesses d'OpenAI, ne serait-ce que celle d'analyser les émotions de ses utilisateurs, sont de fait à prendre avec prudence. « Les expressions du visage correspondent à ce que l'on veut communiquer en société, il n'y a aucun lien avec le ressenti des émotions », avertit par exemple Justine Cassell, directrice de recherche à l'Institut national de recherche en informatique et en automatique et membre du Conseil national du numérique.

L'importance de garde-fous

Pour cette linguiste et psychologue de formation, certaines IA peuvent néanmoins être un appui psychologique efficace en l'absence de professionnels de la santé mentale. A condition qu'elles soient les plus transparentes possible sur ce qu'elles sont capables, ou non, de faire.

« Certaines applications peuvent encourager les gens, à travers l'écrit notamment, à faire sortir leurs émotions et à approfondir leurs réflexions grâce à un système de questions ouvertes, ce qui peut être particulièrement efficace, explique M^{me} Cassell. Mais il faut que le message qu'elles transmettent aux utilisateurs en amont soit clair : "Je peux vous écouter, mais je ne peux pas vous aider véritablement." Dès que l'application propose des solutions et prend la place d'un médecin, c'est là que les dérives commencent. »

D'où l'importance, selon elle, que ces outils ne soient pas prohibés mais plutôt confectionnés avec l'aide de thérapeutes agréés, et répondent à une charte législative. Au Royaume-Uni, un chatbot a ainsi obtenu pour la première fois la certification de dispositif médical fournie par le gouvernement et est utilisé par le National Health Service, le système de santé publique, pour « cibler les patients à risque ». Un modèle similaire existe aux Etats-Unis sous le nom de Wysa et a obtenu une certification de l'agence américaine des médicaments, la Food and Drug Administration.

Les professionnels de santé pourraient eux-mêmes avoir recours à l'IA pour alléger certaines de leurs tâches administratives et ainsi dégager du temps pour prendre en charge davantage de patients. « Certains le font déjà », assure Justine Cassell. Olivier Duris, lui, explique aussi s'en servir depuis plusieurs années comme un outil de médiation avec des enfants autistes. « On peut faire des choses intelligentes avec l'IA, assure-t-il. Mais ça n'a d'intérêt qu'à partir du moment où c'est pensé dans une relation avec un thérapeute, et jamais dans le but de le remplacer. »

