LA MÉDECINE DU FUTUR

INTELLIGENCE ARTIFICIELLE GÉNÉRATIVE ET MODÈLE LINGUISTIQUE LARGE EN SOINS DE SANTÉ.

DU BATTAGE MÉDIATIQUE À LA RÉALITÉ

COUCKE PH A (1)

RÉSUMÉ : Les mots intelligence artificielle générative (GAI) et modèles linguistiques larges (LLM) sont devenus ubiquitaires et incontournables en un temps record. En médecine aussi, le nombre d'applications possibles et preuves de concept foisonnent. Le potentiel est effectivement énorme, en particulier aujourd'hui, dans un monde de soins en déséquilibre, caractérisé par un flux énorme de données et un manque de moyens humains. La force créative de la GAI est illustrée par un exemple issu de la recherche protéomique et le développement de nouveaux médicaments, ainsi que dans le monde de la recherche clinique avec le développement des jumeaux digitaux. Pour le LLM, il faut faire la distinction entre modèles généralistes et spécialisés (adaptés au monde des soins). Là aussi, il devient de plus en plus difficile aujourd'hui d'en faire une revue exhaustive, car la littérature abonde de cas d'usage.

MOTS-CLÉS : *Intelligence artificielle générative - Structure protéique - Jumeaux digitaux - Modèle de langage large - ChatGPT - Aspects éthiques*

GENERATIVE ARTIFICIAL INTELLIGENCE AND LARGE LANGUAGE MODELS IN HEALTHCARE

SUMMARY : Generative artificial intelligence (GAI) and large language models (LLM) made their fulgurant irruption in our society at all levels, inclusive in health care. Possible applications and proofs of concept are everywhere. There is without any reasonable doubt an enormous potential, especially nowadays as we are facing an ever-growing imbalance between the flux of data and the availability of human resources. The creativity of GAI will be highlighted in the quest of defining protein structures and search for new medications, as well as for the development and use of digital twins in clinical research. As far as LLM's are concerned, we need to make a distinction between general models and models dedicated specifically to healthcare. Again, making an extensive overview of used cases is becoming impossible today, facing the numerous publications in the field.

KEYWORDS : *Generative artificial intelligence - Protein structure - Digital twins - Large language models - ChatGPT - Ethical aspects*

INTRODUCTION : DÉFINITIONS UTILES

Tout d'abord, revenons-en aux bases. Pour les néophytes, rappelons que l'IA générative (GAI = Generative Artificial Intelligence) et les modèles linguistiques larges (LLM = Large Language Models) sont fortement liés, le deuxième étant un sous-ensemble du premier.

QU'EST-CE UNE «GAI» ?

Ce terme est apparu pour la toute première fois en 2014, sous la plume d'un jeune scientifique Ian Goodfellow (1). Si on interroge ChatGPT-3.5 à ce sujet, il nous cite les éléments clés qui permettent de définir cette GAI : ce sont des modèles qui apprennent à générer des données et du contenu (textes, images, sons, etc.) de façon créative et innovante, en se basant sur des données d'entraînement. On parle d'«apprentissage automatique» qui fait appel à des réseaux de neurones génératifs ou récurrents. Le terme neurone nous ramène à une notion connexionniste, inspirée par la structure des neurones humains : une entrée de l'information par la «dendrite», la transformation de l'information en signal par le «noyau cellulaire»,

et finalement une sortie du signal par «l'axone». L'algorithme mathématique au niveau du noyau décide d'un seuil d'activation et de la pondération finale du signal.

Quand on met deux réseaux de neurones en compétition (le générateur et le discriminateur), on parle de GAN (Generative Adversarial Network). Ce faisant, on vise à améliorer la capacité du générateur à créer des données que l'on ne pourra pas discerner des données réelles. On retrouve pas mal d'exemples de la puissance de ces GAN dans le monde de l'art, de la musique et de l'écriture. Les plus connus sont DALL-E (OpenAI, San Francisco, CA-USA) et MusicLM (Google, Menlo Park, CA-USA) (2, 3).

COMMENT DÉFINIR UN «MODÈLE LINGUISTIQUE LARGE» ?

Il s'agit généralement d'un modèle de traitement de langage naturel (NLP = «Natural Language Processing»), développé pour comprendre et générer du langage humain. C'est donc *sensu stricto* une forme de GAI, mais limité uniquement au langage. Ce qui caractérise le LLM, c'est un nombre élevé de paramètres, indicatif de sa capacité à capturer des modèles complexes dans les données linguistiques, et ses capacités étendues (comme la compréhens-

(1) Service de Radiothérapie, CHU Liège, Belgique.

sion d'un langage, la génération de texte, la réponse à des questions, etc.). Si, au départ, ils sont entraînés sur de grandes quantités de données, ils sont souvent, par la suite, ajustés à des tâches spécifiques pour un domaine particulier, comme par exemple les soins de santé.

Ce concept de LLM connaît un engouement sans précédent au niveau de la société. Prenons l'exemple de GPT («Generative Pretrained Transformer»), développé par OpenAI. Au départ, le modèle GPT-2 avait déjà impressionné par sa capacité à générer du texte cohérent et pertinent pour un contexte donné et ce, même s'il n'avait été entraîné «que» sur 1,5 milliards de paramètres. Une interface de programmation d'application (API) y a été ajoutée, ce qui permet d'intégrer la version 3 dans diverses applications et services.

Au tournant de la deuxième décennie de ce siècle (2020-2021), ChatGPT est lancé *urbi et orbi* et connaît un succès foudroyant au niveau du grand public. Cette itération de GPT est destinée à un cas d'usage bien particulier : générer du texte conversationnel et répondre à des requêtes dans un contexte de chat. Grâce aux interactions du grand public avec le système, OpenAI a continué à affiner et améliorer cette itération de ChatGPT, à corriger les lacunes et à chasser les incohérences. Aujourd'hui, nous en sommes à la quatrième itération (payante), pour laquelle il faut se préinscrire, tellement la demande est grande.

QUEL EST LE POTENTIEL DE L'IA GÉNÉRATIVE DANS LE SECTEUR DES SOINS ?

L'IA générative a de multiples applications possibles dans le domaine de la santé. De façon générale, la GAI peut servir à générer des données médicales synthétiques (y compris en imagerie, où elle peut aussi servir à améliorer la qualité même de l'image) et ce, pour entraîner les modèles d'apprentissage automatique. Elle démontre son utilité pour la découverte de nouveaux médicaments, pour la simulation d'environnements médicaux réalistes (à des fins d'apprentissage et de formation des professionnels de santé). Elle ouvre la voie vers une approche plus personnalisée et individualisée des traitements (médecine de précision). Elle est destinée à faciliter certaines décisions thérapeutiques. Textes, rapports médicaux, voire même articles scientifiques seront facilement écrits avec son aide. Ce dernier aspect oblige les éditeurs, dès à présent, à établir de nouvelles règles claires pour les auteurs, à exiger

la transparence quant à l'utilisation de la GAI pour la rédaction d'un article scientifique (4). On peut d'ailleurs déjà trouver des articles publiés dans des journaux scientifiques internationaux, dont un des co-auteurs est ChatGPT (5, 6). Ce n'est pas vraiment étonnant puisque, en mai 2023, Tim Boucher, auteur de science-fiction, déclare avoir écrit 97 livres en utilisant ChatGPT et Claude (le LLM d'Anthropic). En plus, il les a illustrés avec des images créées par Midjourney (San Francisco, CA-USA) (7). Ces livres sont toutefois des œuvres assez ésotériques, qui ne suscitent pas beaucoup d'intérêt auprès du lecteur. Par contre, inutile d'expliquer l'engouement des postdocs pour de tels outils, quand il s'agit d'y trouver une aide bienvenue en matière de recherche de littérature, de rédaction d'articles, de thèses et de manuscrits (8).

Nous allons - de façon non exhaustive - illustrer l'importance de la GAI dans deux domaines : le développement de nouveaux médicaments, en utilisant la puissance algorithmique pour la prédiction de la forme des protéines cibles, et l'utilisation du concept du jumeau digital pour la simulation - par exemple - de l'évolution sous traitement et la conduite d'essais cliniques *in silico*.

DE LA PRÉDICTION DE LA STRUCTURE D'UNE PROTÉINE AU DÉVELOPPEMENT DE NOUVELLES PROTÉINES ET LE DÉVELOPPEMENT D'APPROCHES MÉDICAMENTEUSES INNOVANTES

La GAI a indéniablement du potentiel dans le domaine de la recherche. Il y a quelques années, en 2020, Google Deepmind, par le biais de son logiciel AlphaFold, a mis fin à un énorme challenge pour les chercheurs. Ce challenge perdurait depuis des dizaines d'années : la prédiction de la forme des protéines. L'annonce de son efficacité a été faite lors de CASP14 en 2020 (concours organisé depuis 1994 et qui a lieu tous les deux ans; CASP = «Critical Assessment of Structure Prediction») (9). Un an plus tard, Deepmind annonçait avoir cartographié 98,5 % des protéines existantes dans le corps humain.

En 2022, Deepmind n'a pas pris part au CASP15. La compétition s'est d'ailleurs corsée, car non seulement il faut prédire la forme structurale des complexes protéiques avec précision (et pas seulement d'une seule protéine), mais aussi expliquer les interactions possibles avec d'autres molécules comme les médicaments. Un nouveau défi est, par ailleurs, énoncé lors de CASP15 : la prédiction du changement conformationnel de la protéine en fonction d'une mutation.

Si Deepmind est absent de la compétition, son ombre y plane résolument, car les challengers utilisent en grande majorité des méthodes

issues et adaptées d'AlphaFold (10). Par ailleurs, la nouvelle version (AlphaFold2) permet de prédire les différents sous-états conformationnels qui expliquent la fonction biologique de ladite protéine, et qui ouvrent donc la porte pour comprendre et prédire l'effet d'une mutation (11).

En 2023, une forme de GAI appelée Chroma, produit de la société Generate Biomedicines (Somerville, MA-USA), crée des nouvelles protéines *de novo*, qui n'existent pas dans la nature, avec en plus des propriétés thérapeutiques programmables (liaison protéique, contrôle allostérique, activité enzymatique, tous ces éléments qui en définissent le potentiel thérapeutique) (12). Ces formes de GAI fonctionnent un peu comme DALL-E, un générateur d'images (logiciel capable de créer une image à partir d'une commande en langage naturel; le nom DALL-E est inspiré par le robot de Pixar WALL-E et Salvador Dali) : il suffit de donner des instructions comparables à celles que l'on donnerait à DALL-E quand on lui demande de créer une image complètement surréaliste ou loufoque. Ce n'est d'ailleurs pas la seule GAI disponible pour ce genre d'application : le groupe de l'IPD («Institute for Protein Design», Washington, Wa-USA) a publié son approche probante en utilisant RoseTTAFold structure prediction network (13).

Incontestablement, cette application de l'IA, sous toutes ses formes (modèles génératifs, transformateurs et autres), ouvre des portes vers la conception de nouveaux traitements. Certaines molécules issues de cette approche faisant appel à la GAI sont déjà en phase d'essai clinique (14). Mieux encore, la GAI permet de déterminer la probabilité de succès d'un essai clinique et le niveau de précision, selon certains auteurs, atteint 79 % (15). Inutile de préciser que ceci risque bien de changer fondamentalement la recherche préclinique et clinique.

GAI ET JUMEUX DIGITAUX (DIGITAL TWINS)

On parle de plus en plus de l'intérêt d'utiliser des jumeaux digitaux, et la GAI en est la clé de voûte (16). Le concept de jumeau digital est d'application dans pratiquement tous les domaines d'activité humaine. Dans l'industrie, on a vu son apparition dès 2002. Il a pour but, par la création d'un modèle virtuel, de mesurer l'impact de certaines décisions, d'évaluer les risques, de tester de nouvelles approches avant de les implémenter dans le monde réel (17). Dans le secteur médical, son apparition est nettement plus récente. On définit le jumeau digital comme une présentation virtuelle d'un patient, généré par de multiples données du patient lui-même (en utilisant de multiples sources), combinées à des

données de population (phénotypiques et imageries), et des mises à jour en temps réel des données issues des capteurs sur le patient et dans l'environnement (18). L'approche *in silico* permet de simuler l'évolution, par exemple, de la maladie ou de la réponse à un traitement, en ayant apparié le patient réel à son jumeau digital qui, lui, représente les caractéristiques moyennes du groupe le plus proche du patient dans la population (mêmes caractéristiques, y compris les variables cliniques et pronostiques). Le flux de données est bidirectionnel : du patient vers le jumeau digital pour faire des simulations (le principe du «What if») et inversement, du jumeau digital vers le patient pour prendre, par exemple, une décision thérapeutique. Si, au niveau du jumeau digital, on observe la probabilité d'une manifestation indésirable, on peut imaginer un suivi plus rapproché afin d'anticiper cet effet secondaire et intervenir au plus vite en amont des premiers symptômes.

L'intelligence artificielle générative est un moyen puissant pour réaliser ces jumeaux digitaux, car son potentiel de création de nouvelles données, réalistes, complexes et dotées des propriétés souhaitées, est quasiment inépuisable (19). Nous n'en sommes qu'aux prémises, mais on devine déjà son intérêt, en particulier dans le domaine de l'industrie pharmaceutique, des essais cliniques et de l'individualisation de la prise en charge. L'intérêt de l'industrie est d'ailleurs grandissant, ce qui se traduit en nombre de brevets déposés : 73 % sont issus de l'industrie pharmaceutique, et 27 % des universités (18).

QUEL EST LE POTENTIEL DE CHATGPT ET AUTRES MODÈLES DANS LE SECTEUR DES SOINS ?

MODÈLE GÉNÉRALISTE DE LLM VERSUS MODÈLE SPÉCIALISÉ

ChatGPT est un LLM parmi d'autres. En la matière, c'est un peu la course en avant, car chaque acteur de la nouvelle technologie annonce avec fracas la création de son propre modèle. Il y a bien entendu les modèles très populaires comme ChatGPT (OpenAI), mais aussi celui - plus discret - d'Aleph Alpha (start-up allemande qui se positionne comme une alternative européenne au produit d'OpenAI). Les chinois (Baidu et Alibaba) ne sont pas en reste : ils annoncent après Microsoft (Bing), Google (Anthropic et Bard) et Facebook/Meta (Llama), la disponibilité prochaine de leur propre solution. Apple, qui était vraiment à la traîne dans ce

domaine, a décidé d'investir 1 milliard de dollars par an pour rentrer dans la course. Même les géants du commerce en ligne (Amazon) disent développer un LLM, avec deux fois plus de paramètres en comparaison à ChatGPT-4.

À côté de ces modèles généraux issus du monde des GAFA's, il y a des produits plus adaptés en santé. DocsGPT (développé à partir de ChatGPT) ou encore le produit de Hippocratic (start-up à Palo-Alto, CA-USA) sont des exemples plus orientés vers le monde des soins. Le premier annonce qu'il développe une solution qui vise à réduire la charge administrative des soignants, à l'origine d'ailleurs d'une réelle épidémie de fatigue morale dans le secteur. Le second espère bien, avec sa solution, améliorer l'accès aux soins, réduire les coûts de la prise en charge et la fatigue morale. Il déclare avoir dépassé les résultats de ChatGPT-4 sur une multitude de tests et d'exams de connaissance et de certification (105 sur 114, dont 74 avec une différence de 5 % et 43 avec une différence de 10 %, par rapport à ChatGPT).

Le géant Google (Deepmind) développe MedPaLM (Pathway Language Model), un chatbot inspiré par ChatGPT, mais complètement orienté soins de santé contrairement à sa solution Bard plus généraliste (20). PaLM est disponible en source ouverte. Ce modèle PaLM est basé sur 540 milliards de paramètres. Google admet que le modèle n'est pas encore parfait. PaLM est actuellement encore dépassé par l'humain, que ce soit en matière de récupération d'informations, de raisonnement et de compréhension, mais il évolue très rapidement et dépasserait largement tous les autres algorithmes. Google argumente que - très prochainement - le manque d'accessibilité aux soins deviendra un risque plus important que celui représenté par les éventuelles erreurs de leur LLM. La version PaLM2 a été présentée récemment. Cette itération semble - par exemple - capable de décrire des images radiologiques, pour autant qu'on lui pose la question de façon adéquate (21). Il faut donc acquérir une certaine dextérité en matière de «prompts» (manière dont on formule une demande adressée à un LLM). Pas étonnant, dès lors, que certains hôpitaux engagent, dès à présent - et à prix d'or - des ingénieurs spécialisés (AI Prompt Engineer). On se pose, par ailleurs, aussi la question s'il ne serait pas grand temps d'introduire une formation dans les filiales de médecine, afin de préparer les étudiants à l'utilisation adéquate de ce type d'application.

À partir d'avril 2023, PaLM2 a été mis à disposition d'un groupe restreint d'utilisateurs (dont la Mayo Clinic, et le développeur de dossiers médicaux informatisés - DMI - Meditech), afin

qu'il soit mis à l'épreuve (21). Dès la fin de l'été, ils ont élargi le nombre d'utilisateurs. Ils collaborent avec Augmedix (San Francisco, CA-USA) qui veut utiliser le LLM dans le but de créer un système d'écoute ambiante («Ambient Listening System»), capable de transcrire automatiquement dans le DMI les conversations entre médecin et patient. En parallèle, un des leaders mondiaux en matière de DMI, la société EPIC (Madison, WI-USA) déclare étendre sa collaboration avec Microsoft (Redmond, WA-USA) afin de déployer des douzaines de solutions pour, entre autres apports, résumer les notes cliniques, établir le codage des maladies, ou fournir des outils d'exploration de données.

Ce n'est qu'un début : on verra de plus en plus l'invasion des LLM dans le monde des soins. Ces solutions seront tellement intégrées dans les DMI qu'elles passeront quasiment inaperçues pour les utilisateurs. Elles offrent surtout une solution au manque critique de données structurées dans les DMI, à la pénurie toujours croissante de main-d'œuvre médicale, et à la pandémie de fatigue morale dans le secteur.

QUELQUES CAS D'USAGE

Moons et l'équipe de la KUL (Katholieke Universiteit Leuven) ont utilisé et comparé l'efficacité de ChatGPT (OpenAI) et Bard (Google) pour rédiger des textes, afin d'en faciliter la compréhension pour les patients (22). Il y a visiblement des limites à une telle approche, du moins aujourd'hui. Ces auteurs ont effectivement rapporté une tendance à sursimplifier le texte, en occultant une partie importante du contenu. Toutefois, ils s'accordent à dire qu'il y a indéniablement un intérêt à utiliser un LLM pour ce genre d'exercice. Ils encouragent les développeurs de matériel didactique destiné aux patients à faire appel - sous contrôle - à un LLM.

Dans un contexte similaire, une autre équipe s'est posé la question de savoir si ChatGPT augmente ou réduit la littérature médicale en matière de rhinite allergique. À nouveau la prudence s'impose mais, dans l'état actuel, les chercheurs concluent qu'il y a une utilité à faire appel à un LLM comme complément d'information, sans pour autant éliminer - pour l'instant - l'avis médical de l'équation (23). Dans le domaine de la prévention et du screening pour le cancer du sein, les avis donnés par ChatGPT semblent tout à fait corrects, selon une étude conduite par les chercheurs de l'Université du Maryland à Baltimore (24).

On a testé les diagnostics effectués par ChatGPT aux urgences d'un hôpital aux Pays-Bas (Jeroen Bosch Hospital) (25). Il est

important de noter qu'il s'agit d'une étude rétrospective, sur un nombre limité de cas ($n = 30$). Le LLM a été nourri avec les observations cliniques faites aux urgences, la liste des médicaments, et les paramètres biologiques (pris de façon identique et standardisés aux urgences). Chaque cas a été présenté à trois reprises à ChatGPT, afin de contrôler la consistance du diagnostic différentiel établi. Comme pour les autres cas d'usage, on observe une capacité d'accélérer le processus, sans pour autant effacer la nécessité d'un jugement humain expérimenté.

En oncologie, les chercheurs du Mass General Brigham (Boston, MA-USA) ont comparé les avis thérapeutiques donnés par ChatGPT avec les directives édictées par le National Comprehensive Cancer Network en 2021 (NCCN). L'année butoir est importante car les «connaissances» du ChatBot s'étendaient jusqu'en septembre de ladite année. Cela représente forcément une limitation technique, puisque ces modèles LLM évoluent continuellement, et les conclusions ne sont donc plus valables au-delà de la date arrêtée. Ils ont observé une discordance, au moins partielle pour un tiers des cas (cancer du sein, prostate et poumon) (26).

De façon plus large, ChatGPT offre un diagnostic exact dans 77 % des cas, et est adéquat en matière de décision clinique à travers le parcours de soins et ce, pour plusieurs vignettes cliniques issues de différentes spécialités (27). Pour tester cela, l'équipe du Mass General Brigham (Boston, MA-USA) a fait du «prompting» en rafale, agissant un peu comme un médecin artificiel.

Une autre étude a fait appel à 430 participants, recrutés sur une plateforme de crowdfunding pour des études académiques. On a sélectionné une dizaine de questions posées par des patients sur un portail de messageries. Il y a eu systématiquement pour chacune des questions sélectionnées, 5 réponses provenant de soignants et 5 fournies par ChatGPT (le nombre de mots utilisés devait être à peu de choses près le même). On a demandé aux 430 personnes de comparer les réponses afin de déterminer quelles étaient celles fournies par le Chatbot (un genre de test de Turing). Les participants étaient à peine capables de faire la distinction. Cette étude souligne l'importance d'étudier les interactions possibles entre chatbots et patients, car les premiers vont évoluer d'un rôle administratif vers un rôle clinique (28).

En ayant à l'esprit ce rôle sans cesse grandissant des chatbots en clinique, des chercheurs se sont intéressés au niveau de détail et

d'empathie fournis par ceux-ci à des questions posées par des patients (29). Pour cela, on a sélectionné 195 questions provenant de forums publics sur des réseaux sociaux (AskDocs). Les réponses fournies par un médecin et par le chatbot, préalablement «anonymisées» afin qu'il soit impossible de les distinguer, ont été comparées par une équipe de professionnels de la santé, tant au niveau de la qualité de la réponse donnée qu'au niveau de l'empathie témoignée. Pour les experts, les réponses données par le ChatBot sont plus détaillées et font preuve de plus d'empathie.

CONCLUSION

L'avenir de la médecine passe obligatoirement par la case GAN-GAI/LLM, qu'on le veuille ou non. Les opportunités dans le domaine sont énormes. Il devient d'ailleurs impossible de simplement lister toutes les compagnies qui investissent en la matière. On vit très probablement une révolution, et donc l'avènement d'une nouvelle ère en matière de soins.

On observe d'ores et déjà la force créatrice de GAI/GAN dans différents domaines. Nous nous sommes contentés de l'illustrer dans le développement de la recherche expérimentale et clinique. On aurait pu aussi évoquer d'autres cas d'usage comme la génération de données synthétiques, ou encore son application dans la simulation médicale réaliste (pour entraîner les professionnels de soins).

Les LLM, forme restrictive de GAI, ne sont pas en reste. Les applications possibles citées par ChatGPT lui-même sont multiples : assistance aux professionnels de soins (réponse rapide aux questions médicales, aide à la formation médicale) et aux patients (réponse aux questions des patients, soutien émotionnel), automatisation des tâches administratives (prise de rendez-vous, aide à la codification et à la rédaction des rapports médicaux), aide à la prise de décisions cliniques et à la conduite d'essais cliniques.

Il faut toutefois souligner les dangers et les risques de tels développements technologiques. Un LLM généraliste comme ChatGPT a pour objectif d'être convaincant, même si la réponse n'est pas totalement correcte. On parle d'ailleurs d'effet hallucinatoire, tellement le modèle a tendance à exprimer une réponse avec une confiance qui apparaît de prime abord comme illimitée. N'oublions surtout pas qu'il s'agit d'un apprentissage non supervisé, et que les sources de données utilisées ne sont pas toutes - par définition - fiables. Quand il s'agit de ChatGPT,

nous sommes donc en face d'une boîte noire. Par contre, d'autres formes de LLM, spécifiquement développées et destinées au monde des soins, utilisent des algorithmes robustes et explicables, et donc vérifiables. Il est donc impératif d'aborder la technologie en établissant des garde-fous, en mettant en place une réglementation et une gouvernance adéquate, capable, de façon réactive et rapide, de s'adapter à une évolution fulgurante.

Nous nous devons aussi d'aborder - sans tarder - les problèmes éthiques que posent ces développements technologiques accélérés : la discrimination, le risque de stéréotypes et de biais, l'iniquité, la perte de la vie privée, la désinformation, les conséquences sociétales et sociales (sur le marché de l'emploi), les risques écologiques (l'empreinte carbone de par la nécessité d'électricité comme source d'énergie) et la tendance anthropomorphique (qui mène à une confiance exagérée de la part de l'utilisateur) (30, 31).

Ne faisons pas l'autruche : l'utilisation de la GAI et du LLM nous pose indéniablement un problème, en particulier quand il s'agit de la relation de confiance entre soignant et soigné, car les avis peuvent ne pas être concordants. La boîte de Pandore est ouverte : les professionnels du secteur des soins se doivent d'être au courant de l'existence et des limitations de telles applications.

BIBLIOGRAPHIE

- Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks. *arXiv* 2014;1406:2661.
- Chardenon A. OpenAI présente un modèle d'IA capable de générer une image à partir d'une simple phrase. L'Usine Digitale 2021. Dernière consultation en ligne, le 23 novembre 2023. Disponible sur: <http://tinyurl.com/4u5jb2d6>
- Poitiers M. Google lance MusicLM, un outil qui génère de la musique à partir d'un texte. L'Usine Digitale 2023. Dernière consultation en ligne, le 23 novembre 2023. Disponible sur: <http://tinyurl.com/yxnaj9a5>
- Elsevier. The use of generative AI and AI-assisted technologies in writing for Elsevier. Last consulted 2023-11-23. Available from: <http://tinyurl.com/jsf8xj5p>
- O'Connor S. Open artificial intelligence platforms in nursing education: tools for academic progress or abuse? *Nurse Educ in Practice* 2023;66:103537.
- ChatGPT Generative Pretrained Transformer, Zhavoronkov A. Rapamycin in the context of Pascal's Wager: generative pre-trained transformer perspective. *Oncoscience* 2022;9:82-4.
- Nolan B. Sci-fi authors says he wrote 97 books in 9 months using AI tools including ChatGPT and Midjourney. *Business Insider* 2023. Last consulted 2023-11-22. Available from: <http://tinyurl.com/2p93ufkh>
- Nordling L. How ChatGPT is transforming the postdoc experience. *Nature* 2023;622:655-7.
- Callaway E. It will change everything: DeepMind's AI makes gigantic leap in solving protein structures. *Nature* 2020;588:203-4.
- Callaway E. After AlphaFold: protein-folding concept seeks next big breakthrough. *Nature* 2022;613:13-4.
- Wayment-Steele HK, Ojoawo A, Otten R, et al. Predicting multiple conformations via sequence clustering and AlphaFold2. *Nature* 2023. doi.org/10.1038/s41586-023-06832-9
- Ingraham JB, Baranov M, Costello Z, et al. Illuminating protein space with a programmable generative model. *Nature* 2023. doi.org/10.1038/s41586-023-06728-8
- Watson JL, Juergens D, Bennett NR, et al. De novo design of protein structure and function with RFdiffusion. *Nature* 2023;620:1089-110.
- INSILICO MEDICINE. Insilico medicine receives IND approval for novel AI-designed USP1 inhibitor for cancer. *EurekAlert* 2023. Last consulted 2023-11-23. Available from: <https://www.eurekalert.org/news-releases/990417>
- Alliper A, Kudrin R, Bushmakina I, et al. Multimodal AI engine for clinical trial outcome prediction: prospective case study H2 2022 – H2 2023. *ResearchGate* 2022. DOI: 10.13140/RG.2.2.30183.93603
- Emmert-Streib F. What is the role of AI for digital twins? *AI* 2023;4:721-8.
- Tao F, Zhang H, Liu A, Nee A. Digital twins in industry: state-of-the-art. *IEEE Trans Ind Inform* 2019;15:2405-15.
- Coorey G., Figtree GA, Fletcher DF, et al. The health digital twin to tackle cardiovascular disease – a review of an emerging interdisciplinary field. *NPJ Digit Med* 2022;5:126.
- Bordukova M, Makarov N, Rodriguez-Esteban R, et al. Generative artificial intelligence empowers digital twins in drug discovery and clinical trials. *Expert Opin Drug Discov* 2023;27:1-10.
- Singhal K, Azizi S, Tu T, et al. Large language models encode clinical knowledge. *Nature* 2023;620:172-80.
- ICT&Health eds. Google's new AI describes X-rays and answers patient questions. *ICT&Health* 2023. Last consulted 2023-11-23. Available from: <http://tinyurl.com/bdf5pkkb>
- Moons P, Van Bulck L. Using ChatGPT and Google Bard to improve the readability of written patient information: a proof of concept. *Eur J Cardiovasc Nurs* 2023. doi.org/10.1093/eurjcn/zvad087
- Høj S, Thomsen SF, Meteran H, et al. Artificial intelligence and allergic rhinitis: does ChatGPT increase or impair the knowledge. *J Public Health (Oxf)* 2023. Doi: 10.1093/pubmed/fdad219
- Haver HL, Ambinder EB, Bahi M, et al. Appropriateness of breast cancer prevention and screening recommendations provided by ChatGPT. *Radiol* 2023;307:e230424.
- ten Berg H, van Brakel B, van de Wouw L, et al. ChatGPT and generating a differential diagnosis early in an emergency department presentation. *Ann Emerg Med* 2023, in press. doi.org/10.1016/j.annemergmed.2023.08.003
- Chen S, Kann BH, Foote MB, et al. Use of artificial intelligence chatbots for cancer treatment information. *JAMA* 2023;9:1459-62.
- Rao A, Pang M, Kim J, et al. Assessing the utility of ChatGPT throughout the entire clinical workflow: development and usability study. *J Med Internet Res* 2023;25:e48659.
- Nov O, Singh N, Mann D. Putting ChatGPT's medical advice to the (Turing) test: survey study. *JMIR Med Educ* 2023;9:e46939.
- Ayers JW, Poliak A, Dredze M, et al. Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA Intern Med* 2023;183:589-96.
- Mediquality, éditeurs. ChatGPT: entre engouement, controverse et défis éthiques. *Mediquality* 2023. Dernière consultation en ligne, le 22 novembre 2023. Disponible sur: <http://tinyurl.com/4ar9txnp>
- Li H, Moon JT, Purkayastha S, et al. Ethics of large language models in medicine and medical research. *Lancet Digit Health* 2023;5:e333-5.

Les demandes de tirés à part doivent être adressées au Pr Coucke Ph A, Service de Radiothérapie, CHU Liège, Belgique.
Email : pcoucke@chuliege.be