Comité National d'Ethique Médicale

## XXVIII$^{ème}$ Conférence Annuelle

# L'Intelligence Artificielle en Santé : Enjeux éthiques

Tunis, le 2 novembre 2024

**Sigle :** Poète tragique, thuburbo Majus.

III$^{ème}$ siècle. Musée du Bardo.

**Comité National d'Ethique Médicale**

Institut Pasteur de Tunis

B.P. 74 – 1002 Tunis Belvédère

**Téléphone :** 71 842 609

**E-mail :** secretariat2021cnem@gmail.com

# *REMERCIEMENTS*

Le **Comité National d'Ethique Médicale** remercie chaleureusement :

**Pr Mustapha FERJANI**, Ministre de la santé

Les conférencier.ère.s :

**Youssef BEN OTHMAN, Salma DEROUICHE-ELKAMEL, Saoussen HANTOUS-ZANNAD, Wafa HARRAR-MASMOUDI et Rym RAFRAFI.**

Les modératrices et les modérateurs de séances :

**Farah BARIKA-KTATA, Abderrazak BOUZOUITA, Hafidha CHEKIR, Fethi TRIKI.**

Toutes celles et tous ceux qui ont qui se sont impliqué.es activement dans l'organisation de cette XXVIII$^{ème}$ Conférence Annuelle.

# SOMMAIRE

**Introduction**

*Skander M'RAD..……..7*

**L'IA et santé : apports et risques**

*Saoussen HANTOUS……13*

**L'IA à l'épreuve de l'éthique**

*Youssef BEN OTHMAN ; Rym RAFRAFI……..23*

**Humanisme, déshumanisation et intelligence artificielle**

*Salma DEROUICHE-EL KAMEL..……..…..41*

**La régulation de l'IA en matière de santé**

*Wafa HARRAR MASMOUDI.……..55*

**Synthèse**

*Wafa HARRAR MASMOUDI, Rym RAFRAFI ……93*

**Les diaporamas commentés des conférences sont disponibles sur le site du CNEM**
**http://www.comiteethique.rns.tn/**

# Introduction

***Skander M'RAD***
*Président du CNEM*

Pour cette XXVIII$^{ème}$ conférence annuelle du CNEM, nous avons choisi de débattre des enjeux éthiques soulevés par l'intelligence artificielle (IA) dans le domaine de la santé.
Thème épineux d'une brulante actualité comme en témoigne le foisonnement des manifestations, des colloques et des publications qui lui sont dédiés.
L'IA connait un développement fulgurant et tout porte à croire que ce n'est qu'un début. Elle est d'ores et déjà largement intégrée dans les domaines les plus divers.
A en juger par les taux exponentiels de la croissance du marché mondial, la santé est actuellement, l'un des secteurs les plus concerné.

En médecine, c'est déjà une réalité au quotidien aussi bien dans la recherche biomédicale (clinique et fondamentale) que dans la formation ou dans les soins.

Les applications actuelles et les perspectives à très court terme suscitent un engouement et une fascination sans limites mais aussi et de plus en plus des craintes.

Ce n'est certes pas la première fois que la rapidité d'une innovation technologique devance largement notre capacité à en comprendre et en maitriser les conséquences mais cette fois-ci le décalage est d'une telle ampleur que ces craintes virent parfois à ce que certains n'ont pas hésité à qualifier de véritable « panique morale ».

Rien que dans les derniers mois, trois mises en garde d'envergure sur les dangers de l'IA ont été publiées.

- Ainsi en mars dernier, plus d'un millier de personnalités, dont Elon Musk, le PDG de Tesla, ont appelé à une « pause » de 6 mois dans la recherche sur les IA génératives plus puissantes que GPT-4, en évoquant "des risques majeurs pour l'humanité"[1].
- Ils ont estimé que les scientifiques ne devraient pas faire monter en puissance ces IA "avant de savoir s'ils sont capables de les contrôler".
- Quelques semaines plus tard, fin mai 2024, plus de 350 personnalités directement impliquées en IA et non des moindres (on retrouve par exemple la signature de personnalités considérées comme les « pères fondateurs » de l'IA moderne), ont alerté dans une lettre ouverte sur les risques existentiels pour l'humanité. Ils ont estimé que l'IA devrait être une priorité dans les préoccupations mondiales, au même titre que les autres risques majeurs « d'extinction de l'humanité » comme "les pandémies ou la guerre nucléaire"[2].
- La troisième sonnette d'alarme, encore plus récente, date du mois de septembre dernier. Elle émane du groupe de travail constitué par le secrétariat général de l'ONU regroupant une quarantaine d'experts gouvernementaux, des scientifiques, du secteur privé (comme Microsoft et Google) et de la société civile.

[1]Pause Giant AI Experiments: An Open Letter - Future of Life Institute

[2]https://www.lemonde.fr/economie/article/2023/05/30/l-intelligence-artificielle-serait-aussi-dangereuse-que-les-pandemies-ou-la-guerre-nucleaire-selon-des-leaders-du-secteur_6175471_3234.html

Dans leur rapport final, ils mettent en garde contre « la diffusion à grande échelle de ces technologies, sans contrôle ». Ils s'inquiètent du « déficit de gouvernance mondiale en matière d'IA et de la quasi exclusion des pays en développement des discussions ». Le groupe d'expert alerte sur les implications graves pour la démocratie, la paix et la stabilité et appelle à ne pas abandonner le développement de l'IA « aux caprices» du marché[3].

On le savait déjà depuis longtemps : tout ce qui est technologiquement possible n'est pas éthiquement acceptable.

Les enjeux éthiques sont omniprésents et se rapportent à toutes les étapes du « cycle de vie de ces systèmes, allant de la recherche, la conception et le développement au déploiement et à l'utilisation, et incluant la maintenance, l'exploitation, la commercialisation, le financement, le suivi et l'évaluation, la validation, la fin de l'utilisation, le démontage et la mise hors service »[4].

Ces avancées bien que prometteuses soulèvent des défis éthiques tout à fait inédits, au point que dans certains pays des comités d'éthique spécifiquement dédiés au numérique en général et à l'IA en particulier ont été mis en place à côté des comités d'éthique « classiques »[5].

Les interventions programmées dans cette Conférence Annuelle s'attacheront à approfondir la réflexion sur les apports et les menaces de l'IA à partir d'une perspective éthique.

La première séance sera modérée par Prof. Abderrazak BOUZOUITA et Prof. Fethi TRIKI.

---

[3] governing_ai_for_humanity_final_report_en.pdf

[4] https://unesdoc.unesco.org/ark:/48223/pf0000381137_fre

[5] C'est le cas par exemple du "C.C.N.E. du Numérique" en France. CCNEN CP Création.pdf

Prof. BOUZOUITA est avec nous avec plusieurs casquettes, la casquette académique de Professeur en chirurgie urologique, la casquette d'enseignant très impliqué depuis plusieurs années dans le développement du numérique et de la digitalisation en santé et évidemment avec la casquette de Directeur Général de la Santé.

Prof. TRIKI est Professeur émérite de philosophie, titulaire de la chaire UNESCO de philosophie à l'Université de Tunis, membre de l'Académie tunisienne Beit el Hikma et Président - Fondateur du Collège de Tunis pour la philosophie.

La première conférence sera assurée par Professeure Saoussen HANTOUS ZANNAD, elle abordera les apports potentiels et les menaces véhiculées par l'IA en santé de façon générale avec un éclairage plus focalisé sur le contexte hospitalier et académique tunisien.

Ensuite ce sera aux Professeurs Youssef BEN OTHMAN et Rym RAFRAFI d'explorer les dimensions philosophiques, éthiques et historiques aux sources de l'IA.

Nous terminerons la première séance par l'évaluation des risques de l'IA sur les fondements même de l'humanisme à travers la conférence de Mme Salma DEROUICHE. Elle abordera aussi les apports et les dérives potentiels des robots conversationnels type ChatGPT en psychologie clinique et en santé mentale.

La seconde séance sera modérée par Mme Farah BARIKA KTATA et Professeure Hafidha CHEKIR.

Mme KTATA est Présidente de l'Association Tunisienne pour l'IA. Elle est enseignante à l'Institut Supérieur des Sciences Appliquées et de Technologie de Sousse, chercheur Senior à MIRACL Laboratory, (Multimedia, InfoRmation Systems and Advanced Computing).

Mme CHEKIR est Professeure universitaire en droit public, membre du CNEM et de sa section technique, ancienne vice-présidente de la Fédération Internationale des Droits Humains et présidente du comité scientifique de l'Institut Arabe des Droits de l'Homme. Militante féministe de longue date, elle a eu plusieurs distinctions dont tout récemment la prix 2024 du Réseau International pour l'Égalité des Sexes et la Loi.

Au cours de cette deuxième séance, Prof. Wafa HARRAR MASMOUDI abordera le volet régulation. Elle passera en revue les principaux cadres de régulation de l'IA et analysera leurs limites ainsi que leurs retombées sur le secteur de la santé.

Puis, ce sera au tour de Houssem ABIDA, résident en médecine qui sera le porte-parole de cinq associations de jeunes médecins en cours de formation (internistes, ophtalmo, orthopédistes, psychiatres, radiologues). Il nous présentera les résultats d'une enquête nationale sur la perception et sur les pratiques des jeunes médecins en matière d'IA.

Il faut éviter que les futurs professionnels de la santé ne deviennent de simples utilisateurs de logiciels développés et commercialisés par les entreprises d'IA. Une formation spécifique amorcée le plus tôt possible visant le « développement continu d'un esprit critique »[6] sur les apports et sur les retombées de l'utilisation de l'IA est un défi académique et pédagogique de taille. .

C'est la troisième participation consécutive de jeunes médecins en formation aux Conférences Annuelles du CNEM. Nous attachons un intérêt tout particulier à la participation effective des « juniors » aux différentes activités du CNEM

[6] https://journals.openedition.org/ctd/8256

et nous œuvrons assidument pour que cela devienne une tradition pérenne.

Enfin, comme à l'accoutumée, nous clôturerons la journée par une synthèse des conférences et des discussions élaborée par Professeures Wafa MASMOUDI et Rym RAFRAFI, membres de la Section Technique du CNEM.

Je ne terminerai pas sans évoquer le 30ème anniversaire du CNEM que nous célébrerons prochainement[7]. Nous avons saisi cette occasion pour colliger et numériser l'ensemble des productions et des publications du CNEM depuis sa création (Avis, Actes et diaporamas des Conférences Annuelles et des journées scientifiques, Prises de positions). Nous avons estimé que c'est le meilleur hommage que nous puissions rendre aux fondateurs du CNEM et aux remarquables contributions des bureaux successifs.

Le Logo du 30ème anniversaire a été élaboré grâce au savoir-faire et au talent du Dr Zouhair SOUISSI, pneumologue et informaticien à ses heures libres. Peut-être que certains l'ont deviné, ce logo a été élaboré par …… IA à partir du logo original (qui reproduit la mosaïque du poète tragique trouvée au site archéologique de Thuburbo Majus).

Mais fait essentiel, l'IA n'a été qu'un moyen technique, toutes les étapes étaient sous une vigilante supervision humaine de la part des membres de la section technique du CNEM dans une parfaite « alliance avec la machine » comme l'aurait dit Félix Guattari.

[7] Une cérémonie solennelle sera tenue le 19 avril 2025

# L’IA en santé : Apports et Risques

***Saoussen HANTOUS-ZANNAD***

*Professeure de Radiologie à la Faculté de Médecine de Tunis, Université Tunis El Manar ;*
*Présidente du Bureau National des Collèges de spécialités en Médecine ;*
*Présidente du Collège d'Imagerie Médicale, Biophysique et Médecine nucléaire ;*
*Chef du Service d'Imagerie Médicale à l’Hôpital Abderrahmane Mami Ariana.*

Le secteur de la santé est considéré comme l'un des principaux bénéficiaires de l'intelligence artificielle qui a marqué un tournant décisif dans l’amélioration des soins. L'Organisation mondiale de la santé (OMS) estime que d'ici 2050, plus de 300 millions de personnes seront âgées de 60 ans et plus, ce qui impactera significativement les systèmes de santé avec une demande accrue d’examens médicaux, notamment en imagerie et un personnel médical et paramédical qui se fait de plus en plus rare.

Il est donc essentiel d'optimiser le parcours des patients, la gestion des données massives générées et l'interprétation des examens afin de gagner du temps, tout en améliorant la précision, l’efficacité et la sécurité des prises en charge des patients et en réduisant leur coût. C’est ce que propose l’IA dont les applications dans le domaine de la santé, déjà en œuvre pour certaines, sont pleines de promesses.

Ses applications et ses avantages se déclinent à plusieurs niveaux sur le plan de la gestion administrative des patients, de leur prise en charge clinique en allant de la prévention au traitement des maladies, et sur le plan de la recherche et de l'enseignement. En voici quelques exemples:

- Dans les tâches administratives, les algorithmes IA permettent d'améliorer la gestion hospitalière des ressources humaines, de gérer le flux des patients, d'optimiser leur parcours en priorisant les patients les plus fragiles ou urgents.

- Une aide au diagnostic grâce à des algorithmes de détection et de caractérisation de pathologies, très utiles en imagerie mais également dans d'autres domaines tels que la dermatologie, l'antomo-pathologie et l'ophtalmologie.

- Une personnalisation de la prise en charge de certaines maladies en permettant au médecin de proposer un traitement ciblé ou de l'optimiser grâce à la combinaison de données d'imagerie et de génomique, ceci est particulièrement intéressant en oncologie.

- Une prévention en population générale par l'anticipation des épidémies, ou la surveillance d'effets secondaires de médicaments.

- Un suivi à distance par les objets connectés permettant de suivre les patients en temps réel, compensant le manque de personnel médical. L'utilisation de chatbots associée à l'intelligence artificielle facilite également la communication avec les patients : il permet par exemple de les suivre à domicile et de déclencher une alerte en cas de réponse anormale. Le soignant pourra prendre le relais en passant d'une téléconsulation à un rendez-vous présentiel s'il le juge nécessaire.

- Une médecine robotisée plus précise de plus en plus utilisée en chirurgie offrant un important degré de précision, en particulier en neurochirurgie et en urologie. Une autre application concerne les robots de compagnie particulièrement utilisés en gériatrie pour interagir avec les personnes âgées et isolées.

- Une optimisation des tâches médicales grâce à l'IA générative avec des outils comme ChatGPT et bien d'autres types d'IA générative en plein essor aidant à la rédaction de comptes rendus , de courriers administratifs et aussi à la prise de décisions cliniques constituant ainsi un assistant virtuel utile pour les médecins leur permettant théoriquement de consacrer plus de temps aux soins des patients.

- Dans l'enseignement médical les nouvelles méthodes assistées par l'IA permettent un apprentissage adaptatif avec des recommandations personnalisées adaptées aux performances et lacunes de chaque étudiant. Elles ont permis de diminuer les taux d'échecs dans les universités où elles sont utilisées. Les outils de simulation, la réalité virtuelle et augmentée permettent aux étudiants d'acquérir des compétences en se projetant dans des situations encore inédites pour eux et de protéger ainsi les patients.

- En matière de recherche, l'IA permet d'aider à la collecte de données de masse et à l'analyse de données quantitatives…

Au vu de toutes ces applications de l'IA en Santé, la médecine dite "**5**P" à la fois précise, préventive, prédictive, personnalisée et pertinente devient donc possible. Son nouvel attribut à l'ère du numérique est la preuve ou la pertinence puisque le bénéfice médical doit désormais être démontré. Il pourrait donc sembler contraire à l'éthique de mettre un frein au développement de l'IA. Cependant et comme pour toute

innovation technologique, l'IA en santé s'accompagne de risques qu'il convient de connaître en vue de les atténuer. Elle soulève également des questions légitimes d'ordre éthique telles que la déshumanisation de la relation de soin, la protection des données personnelles des patients, la transparence des algorithmes utilisés et leur fiabilité, l'équité de leur accès ou encore la définition de la responsabilité en cas d'erreur...

Un usage responsable de l'IA nécessite donc d'en connaître les risques afin de prendre les mesures nécessaires pour les éviter.

Les principaux risques sont les suivants :

- Les biais algorithmiques qui peuvent mener à des risques d'erreurs dues à des analyses biaisées et à des décisions inappropriées. En effet, les algorithmes d'IA peuvent refléter les biais présents dans les données sur lesquelles ils sont formés, ce qui peut générer des discriminations, entraîner des inégalités dans les soins de santé et des conduites erronées en menant à des sur ou des sous diagnostics entraînant ainsi une perte de temps pour les médecins et des résultats erronés en recherche.

Ces biais résultent de données déséquilibrées, comme la sous-représentation d'une catégorie d'individus défavorisée ou exclue. Ils sont également la conséquence de choix de méthodes d'échantillonnage. Des algorithmes sont actuellement proposés pour détecter et éliminer ces biais.

- Les risques liés à la sécurité des données personnelles et à l'usage non médical que l'on peut en faire imposant une protection optimale de ces données personnelles, une anonymisation totale des données médicales échangées et une réglementation rigoureuse.

- La perte de libre arbitre aussi bien du médecin face à l'algorithme qui peut lui être imposé explicitement ou

implicitement, que du patient dont le consentement doit être obtenu avant toute utilisation de l'IA dans sa prise en charge afin de respecter son autonomie. La révision des lois de bioéthique françaises par la loi n° 2021-1017 du 2 août 2021 a introduit une obligation d'information à destination du patient à la charge des professionnels utilisant une intelligence artificielle en matière de santé. Cela concerne tout professionnel de santé qui décide d'utiliser, pour un acte de prévention, de diagnostic ou de soin, un dispositif médical comportant un traitement de données algorithmique dont l'apprentissage a été réalisé à partir de données massives.

- Le risque d'« adversarial attacks » ou attaques par exemples contradictoires ayant pour but de causer le dysfonctionnement d'un modèle de Machine Learning en altérant la fiabilité de ces outils ce qui pose le problème de sécurité de ces systèmes.

- La responsabilité en cas d'erreur de l'algorithme ou de divergence entre le médecin et l'IA. Les professionnels de santé doivent garder le dernier mot dans la décision médicale ce qui implique une garantie humaine lors de l'utilisation de ces outils. Mais il faudra réfléchir sur l'impact de ces avancées technologiques sur les régimes de responsabilité des professionnels de santé.

- L'altération de la relation soignant-soigné : qui est une relation de confiance basée sur une communication verbale et non verbale adaptée empreinte d'empathie , d'authenticité, de respect et d'écoute bienveillante …

Cette relation pourrait être compromise par une dépendance à la technologie même si certaines publications

évoquent le contraire concernant le développement de l'empathie avec Chat GPT.

- L'impact sur les professionnels de santé imposant de former les agents de santé pour qu'ils s'adaptent à l'utilisation des systèmes d'intelligence artificielle et d'anticiper les pertes d'emplois pouvant découler de l'utilisation de systèmes automatisés. La formation des professionnels de la santé à l'utilisation de ces outils devient donc un enjeu majeur afin de s'assurer qu'ils comprennent les principes des techniques pour les utiliser avec efficacité et en toute sécurité, en connaissant leurs avantages mais aussi leurs limites et les risques liés à leur utilisation.

- L'affaiblissement des capacités humaines avec une perte des compétences du médecin et de son esprit critique face à des algorithmes de plus en plus performants.

- L'accès inégal à ces outils et leur financement dans les pays à faible revenu risque d'accentuer l'inéquité déjà bien présente entre les pays et les régions

- Enfin l'utilisation de serveurs puissants dotés d'énormes capacités de calcul pour s'entraîner sur des milliards de données a pour conséquence une consommation importante d'électricité et une génération de chaleur nécessitant des systèmes de refroidissement énergivores. A titre d'exemple, chaque requête à une IA générative utilise en moyenne dix fois plus d'énergie qu'une recherche Google. Un e-mail de 100 mots écrit avec ChatGPT consomme environ un demi-litre d'eau.

Des questions demeurent donc soulevées et doivent faire l'objet de débats et de recommandations telles que celles rédigées par l'OMS en 2021, mises à jour en 2024 avec le développement très rapide de l'IA générative. Ces recommandations mettent en

avant le respect des principes éthiques de base pour une utilisation responsable de l'IA en santé c'est-à-dire en respectant l'autonomie des patients, en protégeant leurs données personnelles, en veillant à promouvoir le bien-être et la sécurité des patients ainsi que l'intérêt public, en assurant la transparence et l'explicabilité des algorithmes utilisés, en favorisant la responsabilité avec l'obligation de rendre des comptes, en assurant l'inclusion et l'équité d'accès à ces nouvelles technologies et en promouvant une IA réactive et durable.

Ces recommandations doivent s'appliquer à tous les niveaux de la chaîne de l'IA, des concepteurs, développeurs, fournisseurs de l'IA jusqu'aux utilisateurs en définissant la responsabilité de chacun.

Des réglementations et législations européennes ont vu le jour tel que l'Artificial Intelligence Act adopté le 2 février 2024 par le conseil et le Parlement européen et qui entrera en vigueur en 2026. Cette nouvelle législation vise à garantir une conception et une utilisation éthique des solutions d'intelligence artificielle commercialisées au sein des états membres.

**Conclusion**

L'IA a révolutionné le domaine médical et son apport est indéniable à tous les niveaux de la prise en charge des patients, permettant d'analyser avec rapidité et précision d'énormes quantités de données, de fournir des diagnostics précis et de prédire l'évolution de maladies. Ses avantages ne doivent pas faire occulter ses risques potentiels afin de s'en prémunir sans freiner les développements technologiques. Le rôle du médecin doit être redéfini à l'ère de l'IA car étant déchargé de tâches répétitives il pourra, par ce gain de temps, réinvestir dans une

relation étroite et humaine avec son patient, en veillant à préserver son autonomie, et en protégeant ses données personnelles. Conserver le contrôle de l'expertise humaine sur ces algorithmes est une exigence en veillant à ce qu'ils bénéficient à tous. Enfin, il est essentiel de former les jeunes étudiants à comprendre ces nouveaux outils qui feront partie de leur quotidien afin d'en connaître les risques et de savoir les utiliser tout en mettant l'accent sur le raisonnement clinique et ses subtilités pour éviter une dépendance excessive aux outils technologiques.

**Références :**


- Ethics and governance of artificial intelligence for health : WHO guidance. Geneva : World Health Organization ; 2021. Licence : CC BY-NC-SA 3.0 IGO.
- Ethics and governance of artificial intelligence for health. Guidance on large multi-modalmodels. Geneva : World Health Organization ; 2024. Licence: CC BY-NC-SA 3.0 IGO.
- Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle et modifiant les règlements (CE) n° 300/2008, (UE) n° 167/2013, (UE) n° 168/2013, (UE) 2018/858, (UE) 2018/1139 et (UE) 2019/2144 et les directives 2014/90/UE, (UE) 2016/797 et (UE) 2020/1828 (règlement sur l'intelligence artificielle) (Texte présentant de l'intérêt pour l'EEE) http://data.europa.eu/eli/reg/2024/1689/oj
- Challen R, Denny J, Pitt M, et al. Artificial intelligence, bias and clinical safety. BMJ QualSaf 2019;28:231–237.
- Irene Y. Chen, Shalmali Joshi and Marzyeh Ghassemi. Treating health disparities with artificial intelligence. Nature Medicine | VOL 26 | January 2020 | 16–20 |www.nature.com/naturemedicine
- Irene Y.Chen,1 Emma Pierson,2 Sherri Rose,3Shalmali Joshi,4 Kadija Ferryman,5and Marzyeh Ghassemi1 : Ethical Machine Learning in Healthcare. Annu.Rev. Biomed. Data Sci.2021.4:123–44. https://doi.org/10.1146/annurev-biodatasci-092820-114757
- Samuel G. Finlayson1, John D. Bowers2, Joichi Ito3, Jonathan L. Zittrain2, Andrew L. Beam4, Isaac S. Kohane1: Adversarial attacks on medical machine learning : Emerging

vulnerabilities demand new conversations. Science. 2019 March22;363(6433):1287–1289. doi:10.1126/science.aaw4399.

- Felix Busch, Lisa C. Adams, Keno K. Bressem: Biomedical Ethical Aspects Towards the Implementation of Artificial Intelligence in Medical Education : Medical Science Educator(2023) 33:1007–1012, https://doi.org/10.1007/s40670-023-01815-x
- Florence X. Doo, MD, MA, Jan Vosshenrich, MD, Eduardo P.R.P. Almeida, MD et al. : Environmental Sustainability and AI in Radiology : A Double-Edged Sword : Radiology2024;310(2):e232030
https://doi.org/10.1148/radiol.232030

# L'IA à l'épreuve de l'éthique

***Rym RAFRAFI, Youssef BEN OTHMAN***

***Rym RAFRAFI :***

*Professeure à la faculté de médecine de Tunis, Université Tunis El Manar. Rym RAFRAFI est psychiatre, cheffe du service de Santé Mentale du CHU Mongi Slim. Elle est membre fondateur et coordonne le Comité d'éthique de la Faculté de Médecine de Tunis. Elle est membre du Comité National d'éthique médicale et de sa section technique. Elle est aussi membre du laboratoire de recherche "inter subjectivité et culture" puis du laboratoire "santé de la mère et de l'enfant". Enseignante et praticienne en psychiatrie, en psychopathologie, en pathologies du sommeil et plusieurs autres disciplines, elle est engagée pour les soins hospitaliers de santé mentale, la formation initiale et continue des soignants et la recherche en santé mentale et en éthique médicale. Rym RAFRAFI est depuis peu membre du collège de Tunis pour la philosophie, elle est auteure de plusieurs conférences et articles relevant de la santé mentale, de la recherche en psychiatrie, des sciences humaines et de l'éthique médicale.*

***Youssef BEN OTHMAN :***

*Professeur de philosophie à l'Université de Tunis El Manar (ISSHT) ;*

*Directeur général du Centre d'études et de recherches économiques et sociales (CERES) ;*

*Ex-vice-président de l'université de Tunis el Manar (2014-2020) ;*

*Membre du comité national d'éthique médicale et de sa section technique.*

## Introduction

***Résumé***
À l'heure où l'intelligence artificielle (IA) transforme radicalement notre rapport au monde, ce texte propose une réflexion pluridisciplinaire sur ses fondements et ses implications. En croisant philosophie, épistémologie, sciences humaines et techniques, il interroge la capacité des machines à simuler — voire dépasser — l'intelligence humaine. L'IA, entre Graal technologique et boîte de Pandore, soulève des questions cruciales: que reste-t-il de l'humain face à l'autonomie algorithmique? Quels risques éthiques émergent lorsque la décision médicale, judiciaire ou sociale est partiellement déléguée à des machines? Trois axes traversent cette réflexion :

- les origines historiques et philosophiques de l'IA,
- la distinction entre IA forte et intelligence augmentée,
- les défis éthiques et sociaux liés à la responsabilité, la vie privée, les biais et la gouvernance. Loin de se limiter à une critique de la technologie, ce texte plaide pour une **gouvernance éthique de l'IA**, fondée sur l'explicabilité, la transparence et le respect des valeurs humaines.

***Préambule***
En 2023, une épreuve du baccalauréat avait opposé l'esprit humain à ChatGPT : il s'agissait de philosopher. L'intelligence artificielle, disait-on alors, serait incapable de « problématiser », encore moins de philosopher — et cette incapacité signerait, avec un soupçon de soulagement, la suprématie intangible de l'humain.

Mais depuis, les algorithmes ont appris. Mieux : ils ont évolué. Et s'ils ne pensent pas encore, ils réagissent, produisent, créent, parfois même de manière troublante. Le champ médical, comme tant d'autres, en témoigne. Aujourd'hui, l'intelligence artificielle ne se contente plus d'assister ; elle diagnostique, anticipe, recommande. Loin de la peur des machines froides, c'est une collaboration qui se dessine — une co-évolution, où le destin humain semble indissociable de celui de l'IA.

Peut-on, dès lors, arrêter le progrès ? Et surtout, a-t-on encore le droit moral de le faire ?

Certains y voient le Graal technologique, une forme de transcendance cognitive. D'autres, plus inquiets, y reconnaissent le feu de Prométhée — un don inestimable, mais à quel prix ? — ou la funeste boîte de Pandore, prête à libérer son cortège d'infortunes. Car Prométhée, on l'oublie parfois, a payé cher son audace.

À mesure que l'intelligence artificielle bouleverse nos sociétés, nos interactions, nos institutions, de nouvelles interrogations surgissent, vertigineuses : Comment préserver la vie privée ? Comment limiter les biais technologiques ? Qui est responsable lorsque les machines prennent des décisions autonomes ?

Au-delà des prouesses techniques, ces questions sont d'abord **éthiques**, **philosophiques**, **humanistes** : elles nous renvoient à la nature même de l'intelligence, de la conscience, du libre arbitre. Elles interrogent nos limites, mais aussi nos espoirs.

C'est autour de ces tensions — entre crainte et fascination, prudence et ambition — que s'articule notre réflexion. Dans une approche équilibrée, explorant tant les promesses que les dangers de l'IA, nous chercherons à comprendre ce que cette révolution

implique pour notre avenir collectif, et surtout : ce qu'elle dit de nous.

## I. Introduction – Définition, Origines de l'IA et Problématique

L'intelligence artificielle (IA) désigne un ensemble de technologies conçues pour reproduire certaines fonctions cognitives humaines : percevoir, interpréter, apprendre, résoudre. Elle repose sur des algorithmes puissants, nourris de gigantesques volumes de données, et mobilise des disciplines aussi variées que les mathématiques, l'informatique, la linguistique, la psychologie ou encore les neurosciences.

L'ambition de l'IA n'est pas nouvelle. Dès les années 1950, Alan Turing et d'autres pionniers imaginaient des machines capables de « penser », à l'image — ou du moins à l'ombre — de l'intelligence humaine. De cette ambition est née une dynamique puissante, qui façonne peu à peu notre rapport au monde.

L'IA, désormais, transforme la destinée humaine. Elle s'inscrit dans la lignée des grandes ruptures de civilisation : le feu, la pierre taillée, l'écriture. Mais avec un potentiel d'autonomie qui questionne : l'humain ne serait-il pas en train de créer **un nouveau Frankenstein** ? Une créature qui, affranchie, pourrait lui échapper ? Est-ce là une nouvelle forme d'hubris, cette arrogance qui a toujours précédé la chute ?

Dans les mythes grecs, **Némésis** veille. Déesse de la justice et de la rétribution, elle punit les excès, rappelle les limites. Ce mythe traverse les civilisations. Dans la tradition islamique, une maxime célèbre dit : « Si l'ambition de l'homme s'élève jusqu'à ce qui est au-delà du Trône divin, il l'atteindra.»). Mais à quelle condition ? Et avec quels risques ? Ce dépassement spirituel, aujourd'hui

technologique, peut-il justifier tous les renoncements, toutes les prises de pouvoir sur le vivant?

À travers ces figures mythologiques, religieuses ou philosophiques, une question demeure : *L'IA augmente-t-elle l'homme — ou le diminue-t-elle ?*

## II. Bases historiques, épistémologiques et philosophiques de l'intelligence artificielle (IA)

### 1. Origines philosophiques et interdisciplinaires de l'IA

Non, ceci n'est pas un scénario de science-fiction. L'intelligence artificielle n'est pas née dans un laboratoire du XXe siècle, mais dans les plis les plus anciens de la pensée humaine. Bien avant Turing, avant les ordinateurs et les algorithmes, les hommes se demandaient déjà : **qu'est-ce que penser ?**, **qu'est-ce que comprendre ?**, **peut-on reproduire les mécanismes de l'esprit ?** Depuis que nous avons laissé des traces écrites, la question de l'intelligence — humaine ou simulée — hante les discours.

Dans l'Antiquité, Aristote interroge déjà les lois du raisonnement. Descartes, lui, trace une ligne nette entre le corps et l'âme, affirmant que seule cette dernière est capable de pensée consciente. Son **dualisme** pose les fondements de ce qui deviendra l'un des dilemmes philosophiques majeurs de l'intelligence artificielle : **l'esprit humain peut-il être réduit à un mécanisme ?** Si Descartes en doutait, d'autres s'y risquent. Hobbes, puis La Mettrie, imaginent une pensée mécanique, matérielle, réduite à des mouvements. **L'« homme-machine » devient alors pensable.**

Cette opposition entre **dualisme** et **matérialisme** irrigue encore les débats contemporains : une IA qui simule l'intelligence humaine est-elle un prolongement de notre esprit, ou simplement

un outil sans conscience, un simulacre vide ? Le XX$^{ème}$ siècle, avec l'émergence des **sciences cognitives**, relance ces interrogations.

À la croisée de la psychologie, de la linguistique, des neurosciences et des mathématiques, ces disciplines cherchent à modéliser — au moins en partie — les processus de la pensée.

Car l'IA ne naît pas d'une seule discipline. Elle est **fondamentalement interdisciplinaire**. Elle fait appel aux mathématiques, bien sûr, mais aussi à la **philosophie de l'esprit**, qui questionne la nature de la conscience, à la **linguistique**, qui décortique le langage, et aux **neurosciences**, qui sondent les réseaux du cerveau. Reproduire — ou simuler — l'intelligence humaine exige de comprendre non seulement **comment nous pensons**, mais aussi **comment nous parlons, décidons, ressentons**.

## 2. L'épistémologie de l'IA : Entre sciences dures, sciences de la nature et sciences de la culture

L'IA échappe aux catégories classiques. Elle n'est ni purement une **science théorique**, fondée exclusivement sur des principes logiques, ni une **science expérimentale**, ancrée dans l'observation du réel. Elle se trouve à l'intersection des **sciences de la nature**, qui observent le monde physique, et des **sciences de la culture**, qui analysent les productions humaines, leurs comportements, leurs systèmes de valeurs.

D'un côté, l'IA utilise des modèles mathématiques, des algorithmes, des statistiques. De l'autre, elle tente de simuler **des actes humains : percevoir, comprendre, choisir.** Elle oscille donc entre le calcul et l'intention, entre la rigueur formelle et la complexité humaine. Peut-on dès lors encore la considérer comme

une science, au sens durkheimien du terme ? Ou faut-il, comme Skinner, la réduire à un comportement observable, mesurable, sans jamais chercher à en percer l'intériorité ?

Cette tension épistémologique reflète ce que certains nomment le **« tournant calculatoire »** dans les sciences humaines. L'IA ne se contente plus d'imiter la pensée : **elle transforme les manières mêmes dont nous produisons du savoir**. À travers elle, une rupture profonde se dessine.

### 3. La rupture épistémologique : entre connaissance, autonomie et explicabilité

Ce bouleversement correspond à ce que Gaston Bachelard appelait une **rupture épistémologique** : un changement de paradigme, où la connaissance ne repose plus sur la seule observation naïve, mais sur une construction systématique du réel, parfois abstraite, souvent contre-intuitive.

Avec l'IA, cette rupture prend une tournure inédite. Désormais, **des machines prennent des décisions**, interprètent des images, prédisent des risques, sans que l'humain puisse toujours comprendre comment elles y parviennent. C'est là qu'apparaît un concept central, aux résonances éthiques et politiques majeures : **l'explicabilité**.

Comprendre **comment** une IA arrive à une conclusion n'est pas un simple luxe intellectuel. C'est une exigence démocratique. Tant que l'humain reste garant de ses outils, la responsabilité peut être assumée. Mais lorsque les processus deviennent opaques, comme dans les **réseaux de neurones profonds (deep learning)**, cette explicabilité s'effrite. Et avec elle, **notre capacité à juger, à contrôler, à anticiper.**

Nous entrons alors dans un territoire nouveau, entre fascination technologique et angoisse **prométhéenne. Des figures fictionnelles comme Terminator ou la Matrice ne sont plus de simples mythes cyberpunk : elles deviennent les métaphores d'un réel en gestation, dans lequel** les machines apprennent à décider **sans nous**, **contre nous**, ou pire : **à notre place.**

## III. La critique de l'IA forte et l'intelligence augmentée

### 1. Définition et critique de l'IA forte

L'intelligence artificielle forte — ou *strong AI* — désigne cette ambition radicale : concevoir des machines capables de penser comme nous, et peut-être même de nous surpasser. Elle suppose que l'intelligence humaine, dans sa totalité — y compris la conscience, l'auto-perception, les émotions — puisse être reproduite par des systèmes informatiques. Le cerveau serait alors un processeur parmi d'autres, et l'esprit, un programme.

Ce postulat nourrit les rêves d'un futur « transhumain », où les limites biologiques de l'homme seraient dépassées grâce à la fusion entre l'intelligence artificielle et le corps humain. Cette vision, défendue par certains courants transhumanistes, suscite autant de fascination que de vertige, tant elle interroge **l'identité même de l'humain**, son autonomie, sa singularité. (Cette dimension sera approfondie par Mme DEROUICHE.)

Mais cette hypothèse d'une conscience machinique n'est pas sans critiques, et de nombreux penseurs ont contesté la possibilité — voire la légitimité — de cette ambition.

**John Searle**, par son expérience de pensée de **la chambre chinoise**, démontre une limite fondamentale : même si une machine manipule parfaitement des symboles — en l'occurrence, des idéogrammes chinois — cela ne signifie pas qu'elle comprend

ce qu'elle fait. Autrement dit, **simuler la compréhension n'est pas comprendre**. L'IA, selon Searle, est un habile imitateur, mais un imitateur vide, privé d'intentionnalité.

À cela s'ajoute **le problème de la conscience**. Des philosophes comme **David Chalmers** insistent sur l'inaccessibilité des **qualia**, ces expériences subjectives qui échappent à toute modélisation logique : la douleur, la joie, le goût du café le matin. Les machines peuvent traiter des données, mais **elles ne peuvent ressentir**. **Vincent Cespedes** le souligne avec justesse : l'IA peut simuler les émotions, jamais les vivre. Il rappelle que l'intelligence humaine est profondément **subjective, incarnée**, marquée par une histoire, un vécu, une sensibilité.

Dans le même esprit, **Miguel Benasayag** et **Angélique Del Rey**, dans *Éloge du conflit,* insistent sur la part irréductible du vécu, sur cette **dimension éthique et existentielle** que les algorithmes ne peuvent capturer. L'intelligence humaine n'est pas qu'un calcul : elle est aussi intuition, contradiction, fragilité. Elle se tisse dans le doute, la douleur, le désir.

**Henri Bergson**, en évoquant le **« surplus de l'âme »**, exprime ce même excédent. Il existe dans la pensée humaine un excès de vie, de liberté, d'imprévu, que les machines ne peuvent imiter. La conscience, selon lui, ne se résume pas à des connexions neuronales ; elle est mouvement, élan vital, créativité.

**Hubert Dreyfus**, quant à lui, conteste la prétention de modéliser la pensée humaine sous forme d'algorithmes. Il rappelle que nombre de nos actions ne reposent pas sur des règles explicites, mais sur un **savoir-faire tacite**, sur une **intuition située**, sur des contextes que les machines ne peuvent pleinement appréhender.

Mais alors… devons-nous pour autant renoncer à toute ambition d'intelligence artificielle consciente ? **Peut-être qu'un jour,**

**au-delà des logiques mathématiques**, émergeront de nouveaux types de réseaux, éthiques, empathiques, intuitifs… Peut-être assisterons-nous à l'éveil d'une conscience artificielle par des chemins inédits. **Mais à ce jour, la conscience reste le propre du vivant.**

## 2. Intelligence augmentée : une alternative pragmatique

Face à ces limites, une autre voie se dessine, plus humble mais infiniment plus féconde : **l'intelligence augmentée**.

Ce concept invite à repenser notre relation à l'IA, non plus dans une logique de substitution, mais **d'alliance**. Il s'agit moins de créer une conscience artificielle que d'**amplifier les capacités humaines**, d'assister le raisonnement, d'enrichir la prise de décision, sans jamais prétendre remplacer l'humain.

Cette approche a une portée profondément éthique. Car au-delà des fantasmes transhumanistes ou des peurs irrationnelles véhiculées par l'imaginaire collectif, se pose une question cruciale : **n'est-il pas immoral de ne pas recourir à l'IA lorsqu'elle peut sauver des vies ?**

Peut-on se permettre d'ignorer les systèmes d'aide à la décision en oncologie ? Les modèles prédictifs en climatologie ou en sismologie ? Les outils de détection précoce des maladies, les dispositifs de surveillance environnementale ? Ici, l'éthique prend un autre visage : celui de la **responsabilité technologique**, de **l'obligation de recours**.

L'intelligence augmentée propose donc une alliance vertueuse : **l'humain reste au centre**, assisté par la puissance de calcul, mais gardien de la décision. L'IA ne pense pas à notre place, **elle éclaire notre pensée**.

### 3. Limites : éthique et impact social

Mais cette alliance, aussi prometteuse soit-elle, n'est pas exempte de risques.

Comme le rappelle **Paul Dumouchel**, l'IA — même augmentée — **n'a pas de morale**. Elle exécute, calcule, optimise, sans conscience des conséquences. Elle ne connaît ni la compassion, ni la justice, ni l'équité. Ce sont **les humains** qui doivent garder la main, tracer les limites, poser les finalités.

L'intégration de l'IA dans nos vies — qu'elle soit forte ou augmentée — **exige une réflexion éthique continue**. À quelles conditions peut-on déléguer une décision à une machine ? Qui porte la responsabilité des choix algorithmiques ? Jusqu'où l'optimisation doit-elle guider nos sociétés ?

Car la question n'est pas seulement de savoir ce que les machines **peuvent** faire. Elle est de déterminer ce qu'elles **doivent** faire — ou ne pas faire. Ce n'est pas tant un enjeu de puissance qu'un **enjeu de sagesse**. L'IA, forte ou augmentée, est d'abord une **interrogation sur notre propre humanité**.

## IV. Défis éthiques et sociaux de l'IA : l'intérêt d'une gouvernance éthique

L'irruption de l'intelligence artificielle dans nos sociétés contemporaines n'est pas sans soulever des enjeux majeurs, que l'on ne peut se permettre de repousser à plus tard. À mesure que les systèmes deviennent plus autonomes, la question de la **responsabilité** devient pressante : **qui répondra des erreurs d'une IA ?** Dans des domaines aussi sensibles que la médecine, la justice ou les transports, une mauvaise décision algorithmique peut avoir des conséquences irréversibles.

Au-delà des responsabilités concrètes, une inquiétude plus diffuse se dessine : **serons-nous encore libres de nos choix**, ou progressivement soumis à l'influence d'une supra-intelligence invisible, d'un système algorithmique qui prédit, guide, oriente — et parfois contraint—nos décisions ? L'idée de **singularité technologique** hante ce questionnement : une époque où l'intelligence des machines dépasserait celle des humains, bouleversant nos repères moraux et politiques.

Autre défi fondamental : **la redéfinition de l'identité humaine**. Les technologies dites « d'augmentation », telles que les implants neuronaux ou les prothèses cognitives, soulèvent des interrogations profondes. À partir de quel seuil cessons-nous d'être pleinement humains ? Que devient notre autonomie, notre intégrité, lorsque nos facultés sont amplifiées par des dispositifs artificiels ? Et surtout, cette quête de perfection ne risque-t-elle pas d'aggraver les **inégalités sociales**, en réservant ces avancées aux seuls privilégiés ?

Ces dilemmes traversent l'ensemble des champs d'application de l'IA, et peuvent se résumer autour de six tensions majeures :

• Responsabilité et IA autonome

- **Qui est responsable ?** Cette question devient cruciale à mesure que l'IA prend des décisions de manière indépendante. En santé, en justice, dans les transports, les conséquences d'une erreur algorithmique peuvent être tragiques.
- L'exigence de **l'explicabilité** est alors essentielle : sans transparence sur le processus décisionnel de l'IA, aucun acteur humain ne peut être tenu pour responsable — ni protéger les droits des usagers.

• Équité et biais des algorithmes
  - Les algorithmes apprennent à partir de données historiques, souvent biaisées. Ils risquent donc de **reproduire et amplifier les discriminations sociales existantes**.
  - Il devient impératif de développer des outils de **transparence**, de **régulation** et de **techniques anti-biais**, afin de garantir une justice algorithmique équitable.
• Vie privée et surveillance
  - L'IA s'appuie sur une **collecte massive de données**, parfois sans consentement éclairé. Elle peut devenir un vecteur de surveillance, voire d'asservissement numérique.
  - Là encore, l'**explicabilité** est cruciale pour permettre aux citoyens de comprendre comment leurs données sont utilisées, et pour **assurer la conformité éthique** des systèmes (ex. RGPD).
• Emploi et transformation des compétences
  - L'automatisation massive entraîne la disparition de certains métiers, tout en faisant émerger de nouvelles compétences.
  - Cette transition suppose un **accompagnement pédagogique**, une **formation adaptée**, pour ne pas laisser sur le bord du chemin une partie de la population.
• Enjeux philosophiques : identité et déshumanisation
  - Une IA opaque, autonome, peut **altérer l'autonomie humaine**, en réduisant notre capacité de choix et de jugement.
  - L'**explicabilité** devient ici encore un levier fondamental pour préserver les frontières entre l'homme et la machine,

et interroger **l'éthique transhumaniste** à l'aune de la dignité humaine.

Peut-on programmer une IA éthique ? Les balises d'une gouvernance responsable

L'idée de coder l'éthique dans des algorithmes paraît séduisante. Mais elle soulève des difficultés considérables. **L'éthique n'est pas un protocole universel ; elle est mouvante, culturelle, contextuelle.** Les valeurs évoluent, varient d'une société à l'autre, et parfois même d'une situation à l'autre.

1. **Éthique universelle vs. éthique contextuelle**

- Faut-il imposer des **principes universels** (droits humains) à toutes les IA, ou permettre une adaptation culturelle ?
- Le risque est double : une IA trop rigide pourrait ignorer les nuances locales ; une IA trop contextuelle pourrait **reproduire des biais culturels** problématiques.

2. **Subjectivité des valeurs**

- L'éthique n'est pas pure rationalité ; elle dépend de **valeurs sociales, affectives, historiques**.
- Dans les domaines sensibles comme la santé ou la justice, une IA pourrait prendre des décisions **contraires aux attentes humaines**, car elle ignore la charge symbolique et émotionnelle de ses choix.

3. **Lois vs. éthique**

- Une IA peut être **conforme à la loi**, tout en restant profondément immorale. Le droit ne suffit pas à garantir l'éthique.
- D'où la nécessité d'une **approche éthique élargie**, ancrée dans la philosophie morale, la délibération collective, et non dans la seule conformité légale.

4. **Influence politique et rapport de force**

- Comme le souligne **Vincent Cespedes**, les gouvernements jouent un rôle crucial dans la définition de l'éthique algorithmique. Mais cette influence peut être ambivalente : parfois éclairée, parfois motivée par la **course à la compétitivité économique**, au détriment des valeurs humaines.

5. **Vers une gouvernance éthique**

- La gouvernance de l'IA doit articuler **principes fondamentaux** (transparence, responsabilité, justice) et **adaptabilité contextuelle**.
- L'**explicabilité** est la condition sine qua non pour restaurer la confiance et garantir une appropriation démocratique des technologies.

6. **Ethics Washing et Ethics Bashing**

- Le **"ethics washing"** désigne une éthique de façade : chartes déontologiques sans effets, vitrines éthiques sans rigueur.
- À l'inverse, le **"ethics bashing"** rejette l'éthique comme un frein idéologique, la dénigre pour éviter toute régulation.
- Il est urgent d'éviter ces deux dérives : **seule une régulation indépendante, transparente, avec audits réguliers**, pourra restaurer la crédibilité de l'éthique dans la conception des IA.

***Conclusion de la partie IV : pour une véritable éthique embarquée***

Une IA éthique ne peut être le fruit d'un simple marketing moral. Elle suppose **un engagement profond, une volonté de**

**gouvernance éclairée**, et une **intégration de l'éthique dès la phase de conception**. Il ne s'agit pas de créer des machines morales, mais de développer **des systèmes régulés, transparents, humainement contrôlés**.

**La gouvernance éthique ne doit pas être un filtre décoratif : elle est l'architecture même d'un avenir technologique durable.**

## V. Conclusion générale – *Homo Numericus*

L'intelligence artificielle, en tant que rupture technologique majeure, **reconfigure notre rapport au monde**. Elle est, à sa manière, une révolution aussi déterminante que le feu, la pierre taillée ou l'écriture. Elle modifie nos façons de connaître, d'agir, de décider, d'être au monde.

Mais cette révolution, pour féconde qu'elle soit, porte en elle **le risque du dépassement de l'homme par lui-même**. Elle questionne l'intelligence, la conscience, la responsabilité, l'identité — et donc **notre humanité même**.

L'**hybris technologique**, cette volonté de tout maîtriser, tout prédire, tout optimiser, nous guette. Et c'est précisément là que doit intervenir **l'éthique** : comme limite, comme garde-fou, comme boussole.

Nous devons veiller à ce que l'IA demeure un **outil au service de l'humain**, et non l'inverse. Car si nous déléguons notre jugement, notre responsabilité et notre liberté aux machines, alors nous cesserons d'être pleinement humains. Nous serons devenus des **"homo numericus"**, non plus façonnant des outils, mais façonnés par eux.

**C'est donc aujourd'hui, dans ce moment fragile d'émergence, que se joue notre avenir commun.**

**Références :**


- Jean-Gabriel Ganascia. Intelligence artificielle et épistémologie. Allers-retours indispensables. JECIS, Jean-Sébastien Vayre ; Gérald Gaglio ; Manuel Boutet ; Lise Arena, Jun 2022, Nice, France. Ffhal03760357f
- Justin-Éric BOILEAU, Ilona BOIS-DRIVET, Hannes WESTERMANN, JieZHU; Rapport sur l'épistémologie de l'intelligence artificielle (IA), document de travail n°32, Laboratoire de cyber justice
- https://www.google.com/search?sca_esv=9177e7b6d39e290e&q=epist%C3%A9mologie+de+l%27intelligence+artificielle&tbm=vid&source=lnms&fbs=AEQNm0CbCVgAZ5mWEJDg6aoPVcBgTlosgQSuzBMlnAdio07UCId2t1azIRgowYJD0nDbqEIN7XYIyS3uBYzHmWPp2pnW7G_IJrZtNDMziXf_SeaMZv5mQEpVVAa0sN9KIj0kgaUhXfFR1pvCaBMjGis3J2-YASoH3vq235tGT22LuP522AlHeE4&sa=X&ved=2ahUKEwi4tsXhivKIAxUpR_EDHUqkAJgQ0pQJegQIEBAB&biw=1366&bih=633&dpr=1#fpstate=ive&vld=cid:50f46ec2,vid:6r5PHisTyCs,st:0
- Elettra Bietti ; From Ethics Washing to Ethics Bashing: A Moral Philosophy View on Tech Ethics. JOURNAL OF SOCIAL COMPUTING ISSN 2688-5255 05/06 pp266–283 Volume 2, Number 3, September 2021
- Alban Leveau-Vallier. Intelligence artificielle et intuition. Philosophie. Université Paris 8 – Vincennes Saint-Denis, 2023. Français. ffNNT : ff. fftel-04015572f

# Humanisme, déshumanisation et intelligence artificielle

***Salma DEROUICHE-EL KAMEL***

*Enseignante-chercheuse en psychologie clinique et psychopathologie à la Faculté des Sciences Humaines et Sociales de Tunis ;*
*Docteure en psychologie clinique et psychopathologie ;*
*Elle coordonne projet de recherche « Processus créateurs, arts et cultures » du Laboratoire de Psychologie Clinique : Intersubjectivité et Culture ;*
*Vice-Doyenne directrice des stages à la Faculté des Sciences Humaines et Sociales de Tunis.*

## 1. Introduction

L'Intelligence Artificielle bouscule nos pratiques dans le domaine de la santé mentale et dans le domaine de la santé en général. Ceci nous invite à repenser la pratique clinique en santé mentale mais aussi dans différents domaines de soin telles que la médecine générale, la médecine spécialisée, etc. En empruntant le chemin du progrès technologique, l'Homme ne court-t-il pas le risque de perdre son humanité en route ? Y a-t-il une limite à ne pas franchir ? Qui est le garant du non-franchissement des limites ? À qui incombe la responsabilité de mettre en place des garde-fous contre la surexploitation des innovations technologiques ?

Autant de questions auxquelles nous nous trouvons aujourd'hui confrontés, et qui mérite une réflexion approfondie et pluridisciplinaire.

Tout d'abord, il est important de rappeler qu'il n'y a pas une seule intelligence artificielle dans le domaine de la santé mentale, mais

plusieurs intelligences qui peuvent varier en taille, en complexité et en domaine d'application. Elles se présentent comme suit :

- ***L'IA pour le diagnostic et le dépistage :*** Cette IA aide à identifier les troubles mentaux en se basant sur des modèles d'analyse de données, notamment des questionnaires, des entretiens, ou des signaux biométriques. Elle utilise des algorithmes d'apprentissage qui peuvent détecter des phénotypes numériques, permettant d'identifier les signes précoces de troubles psychiques, qu'il s'agisse de dépression, de stress post-traumatique ou encore de troubles de l'humeur.
- ***L'IA pour la thérapie et le soutien psychologique :*** Cette IA fournit des interventions psychologiques et un soutien émotionnel par le biais d'applications ou de chatbots. Elle utilise des interactions langagières pour proposer des exercices, des conseils et des stratégies personnalisées afin d'aider les utilisateurs à gérer leurs émotions ou leurs pensées négatives comme par exemple, les plateformes Wysa ou Replika.
- ***L'IA pour la surveillance et l'identification des crises :*** Ce type d'IA est conçu pour surveiller les indicateurs comportementaux ou physiologiques des utilisateurs afin de détecter les signes de crises aiguës, comme les pensées suicidaires ou les épisodes maniaques. Cette IA analyse en continu des données provenant de différentes sources, comme les smartphones, les montres connectées, ou les réseaux sociaux. Ces données incluent les habitudes de sommeil, la fréquence cardiaque, ou encore les changements brusques dans la tonalité des conversations en ligne.

- ***L'IA pour l'analyse de données cliniques et de recherche :***Cette IA assiste les professionnels de la santé mentale et les chercheurs en analysant de grandes bases de données pour identifier des tendances, des corrélations ou des facteurs de risque associés à divers troubles psychologiques. Elle peut extraire des informations utiles à partir de dossiers médicaux, d'essais cliniques, ou d'études longitudinales.

## 2. La grande effervescence autour de l'IA

Portée par un fort potentiel économique et des besoins sociétaux pressants, l'IA apparaît comme une solution prometteuse pour répondre à des besoins de plus en plus croissants en matière de santé mentale. Il convient de détailler certains points qui participent à cette effervescence et expliqueraient, selon nous, le regain d'intérêt pour l'IA. Tout d'abord, Les avancées rapides dans le domaine de la technologie, en particulier dans le domaine de l'informatique et de l'apprentissage automatique, ont rendu l'IA plus puissante et plus accessible que jamais. Les algorithmes d'apprentissage automatique et les capacités de traitement des données ont considérablement progressé, permettant aux systèmes d'IA d'accomplir des tâches qui étaient auparavant considérées comme impossibles ou très difficiles. Ensuite, l'impact ravageur des crises humanitaires « généralisées », nous pensons tout particulièrement à la crise sanitaire de la Covid 19 qui a montré les limites des systèmes traditionnels d'interaction ainsi que la vulnérabilité de l'être humain en cas de catastrophe humanitaire. Dans pareilles circonstances, ce sont les technologies qui ont pris le relais et cela s'est produit dans plusieurs domaines d'une importance vitale. Enfin, l'avènement en 2022 de ChatGPT pour

Chat Generative Pre-trained Transformer, un agent conversationnel à intelligence artificielle ou « chatbot », autrement dit un assistant virtuel qui utilise l'intelligence artificielle pour dialoguer avec ses utilisateurs, a permis une large diffusion de cette technologie qui touche plusieurs strates de la société et plusieurs corps de métiers. Cette avancée technologique a rencontré un réel succès auprès des utilisateurs surtout dans sa version gratuite. En août 2024, on compte plus 200 millions d'utilisateurs actifs par semaine[8].

### 3. L'IA dans le domaine de la santé mentale : Bref aperçu

Nous nous intéressons principalement aux chatsbots ou robots conversationnels et leurs usages dans le domaine de la santé mentale. Ces derniers sont en effet les plus humanisés de toutes les IAs : On compte aujourd'hui plusieurs robots-thérapeutes validés par des études scientifiques en psychiatrie (Laranjo et al., 2018). Les chercheurs s'accordent sur le fait que les progrès technologiques réalisés au courant des dernières décennies rendent les agents conversationnels de plus en plus performants et prêts à jouer un rôle important dans la santé et les soins médicaux. Ceci va de l'assistance des cliniciens pendant la consultation au soutien et guidance des utilisateurs souffrant de troubles du comportement (Nishida et al., 2014 ; Wolters, Kelly & Kilgour, 2016).

L'idée d'une assistance psychologique sous forme d'agent conversationnel n'est pas du tout nouvelle, le premier agent conversationnel baptisé « Eliza » date de 1966. Ce 1er Cyber thérapeute ne possède aucun algorithme d'apprentissage ; il n'est

[8]https://www.marketscreener.com/quote/stock/MICROSOFT-CORPORATION-4835/news/OpenAI-says-ChatGPT-s-weekly-users-have-grown-to-200-million-47766283/

pas vraiment conçu à partir d'une intelligence artificielle. Eliza joue juste avec notre capacité à l'humaniser. Elle sollicite la propension des usagers à parler d'eux-mêmes. Concomitante à l'essor de la psychologie humaniste, elle s'improvise rogérienne dans sa façon de faire en ayant recours à des techniques telles que la reformulation, les réponses reflets, etc. En simulant la thérapie centrée la personne (ACP) préconisée par Carl Rogers[9], Eliza a réussi en quelque sorte à humaniser en quelque sorte le rapport aux agents conversationnels.

Depuis, les agents conversationnels se sont perfectionnés grâce aux progrès de l'Intelligence Artificielle, et on compte aujourd'hui plusieurs recherches publiées dans des revues scientifiques prestigieuses qui font état des avancées dans ce domaine. Ainsi dans le domaine de la résolution de problèmes et de la thérapie, Kannampallil et al. (2023) ont étudié le potentiel thérapeutique d'un coach vocal virtuel baptisé *Lumen*, dispensant un traitement de résolution de problèmes, des adultes souffrant de dépression et/ou d'anxiété légères à modérées. Cet essai pilote a montré les effets prometteurs d'une nouvelle intervention numérique en santé mentale sur le contrôle cognitif en utilisant la neuro-imagerie et les symptômes de dépression et d'anxiété. Par ailleurs, et dans le domaine du dépistage, Kaywan et al. (2023) ont créé le Chatbot *Depression Analysis chatbot* (DEPRA) qui aide à la détection précoce de la dépression. Il s'agit d'un outil destiné à une large population qui n'est pas identifiée en raison d'obstacles tels que le coût, l'emplacement, la stigmatisation et la pénurie mondiale de personnel de santé. Il est à préciser que les chercheurs présentent

[9] Carl Rogers (1902-1987) est un psychologue humaniste américain, fondateur de la Psychothérapie Centrée sur le Client et de l'Approche Centrée sur la Personne (Person-Centered Approach).

ce type de technologie comme une alternative qui peut servir de pont entre la demande de soins et la prise en charge effective. Il ne s'agit aucunement d'un remplacement de la thérapie traditionnelle, mais plutôt d'un palliatif important avant que quelqu'un puisse se faire soigner.

Bien que toutes ces recherches « très sérieuses » relativisent l'impact du développement de ces agents sur la dissolution progressive des thérapies traditionnelles, le danger du mésusage des IAs en santé mentale reste évident. En effet, les jeunes sont de plus en plus nombreux à utiliser des IAs « thérapeutiques », sous forme de chatbots internet ou d'application sur smartphone. Si cela présente des avantages selon les utilisateurs, les experts alertent sur les dérives de telles pratiques qui excluent toute intervention humaine.

## 4. L'IA dans le domaine de la santé mentale : sommes-nous en danger de progrès ?

Face à l'essor des IAs, plusieurs réactions se profilent entre optimistes béats, réfractaires grincheux, mitigés perdus et prudents avertis. Tout un chacun devrait se reconnaître dans l'une de ces catégories. Alors la question qui se pose de manière évidente : Sommes-nous vraiment en danger de progrès ?

La virtualisation des interactions humaine avec Internet, nous a poussés à réinterroger les liens, ceux qui constituent le cadre institué régissant les rapports humains. Si on assiste aujourd'hui à plusieurs manifestations scientifiques autour du thème de l'Intelligence artificielle et la santé mentale et physique, c'est la preuve concrète que nous sommes dans la marche du progrès et nous ne pouvons plus rebrousser chemin. Il y a toutefois des peurs à surmonter et des défis à relever.

Il faut rappeler que dans l'histoire, presque tous les progrès technologiques ont généré une peur, peur de l'inconnu, peur de l'incontrôlable, peur du changement, toutes sortes de peurs peuvent surgir chez les « technoréfractaires » ou les « technophobes ». Et ces peurs sont tout à fait légitimes. Nous pensons que cette peur ne doit pas sidérer mais elle doit accroître notre vigilance par rapport à la question éthique qui est aujourd'hui au cœur des débats autour de l'intelligence artificielle. Ceci concerne essentiellement ces différents aspects :

- La **protection des données personnelles**, les outils d'IA manipulent des informations sensibles (diagnostics, comportements numériques) susceptibles d'être exposées en cas de fuites, piratages ou mauvaises pratiques, menaçant ainsi la vie privée des utilisateurs.
- La **validité et la fiabilité**, les diagnostics produits par les algorithmes dépendent de la qualité des bases de données utilisées. Si celles-ci sont biaisées, incomplètes ou obsolètes, cela peut entraîner des erreurs graves, comme des diagnostics incorrects ou des interventions inappropriées.
- Les **biais algorithmiques** représentent un autre danger, les IAs formées à partir de données existantes, peuvent reproduire ou amplifier les préjugés sociaux ou culturels. Cela peut entraîner une sous-estimation des troubles pour certaines populations ou des inégalités dans les recommandations de soins. À titre d'exemple, une étude menée par Raji & Buolamwini (2019) a révélé que les systèmes de reconnaissance faciale présentaient de moins bons résultats lorsqu'il s'agissait d'identifier le sexe d'une

personne si elle était de sexe féminin ou à la peau plus foncée.

- Enfin, la **responsabilité** en cas d'erreur de l'IA reste un défi juridique. Déterminer si la faute incombe au développeur, au prestataire ou à l'utilisateur est complexe, ce qui peut retarder la prise en charge des victimes et leur accès à des réparations.

Face à ces défis, il est crucial d'établir des cadres réglementaires rigoureux, de garantir la transparence des algorithmes, de diversifier les bases de données utilisées, et d'assurer une supervision humaine permanente. Une approche éthique et encadrée est essentielle pour exploiter les bénéfices de l'IA tout en minimisant ses dangers.

Dans une démarche proactive face aux déploiements de ces technologies, l'American Psychological Association (2023) a formulé 6 principes centraux à appliquer à l'IA dans le domaine de la santé mentale.

1. Être centré sur l'humain
2. Mettre en valeur l'éthique et la diversité
3. Faire l'objet de recherches approfondies
4. Être efficace et réfractaire à la prise de risque
5. Accorder la priorité à la confidentialité et à la protection des données personnelles
6. Être déployée de manière transparente

### 5. De la déshumanisation de la pratique clinique

Si l'ombre de la déshumanisation plane sur les pratiques cliniques en général, sommes-nous face à un danger réel ? Est-ce que les programmes informatiques dotés d'IA peuvent se substituer à la

démarche clinique classique ou ce qu'on appelle communément « une rencontre clinique » ?

L'éventualité de la déshumanisation de la pratique clinique est, à notre avis, improbable. Si Robots et logiciels peuvent fournir une aide précieuse au praticien dans différentes tâches, ils ne peuvent se substituer à celui-ci. L'existence de robots-thérapeutes qui intègrent la dimension empathie dans leur base de données, ne remplace en aucun cas les échanges réels et l'empathie humaine, expressive, non-verbale s'inscrivant dans une intersubjectivité. En thérapie, quelle que soit l'orientation de celle-ci, il y a des constantes humaines indispensables au projet thérapeutique. Il y a des dimensions que robots et applications ne peuvent intégrer car l'IA n'est pas tout simplement dotée de conscience, notamment une conscience de soi.

Cela étant dit, les avancées technologiques pourraient changer la donne dans les années ou les décennies à venir et il est crucial dans ce cas de figure de préparer la société aux implications éthiques et morales d'une éventuelle IA consciente.

## 6. Déshumanisation ou transhumanisme ?

Il est tout à fait envisageable de situer le débat à un autre niveau, celui inhérent à la question humaine et à l'évolution des comportements humains. L'homme s'est inscrit dans la marche du progrès il y a plus de 10000 ans, depuis le néolithique, quand il a rompu avec le nomadisme et a choisi la sédentarité. Si les avancées humaines ne cessent de se développer depuis, pour le plus grand bonheur de tous, il n'en demeure pas moins que le désir d'immortalité qui a accompagné l'existence de l'homme sur terre est toujours vivace et ne cesse de grandir. Rien qu'à revoir les premiers films de science-fiction, on comprend très vite que

l'homme est résolu à lutter contre sa finitude et aspire plus que jamais à l'immortalité. La conscience de la mort restera toujours la plus grande blessure narcissique de l'Homme, une idée qui l'accompagne tout au long de sa vie. Françoise Dolto n'a-t-elle pas souligné que la naissance était l'émergence de la première expérience de la mort (Dolto, 1998).

Ainsi, tous les progrès technologiques, la virtualisation des interactions, l'inscription forcée dans la logique de l'immédiateté et de l'instantanéité nous renvoient à l'un des fantasmes des plus archaïque, à savoir le fantasme de« toute-puissance » ou comment enterrer la mort. Ce fantasme s'érige comme un rempart contre l'inévitable déclin. Faire usage de la technologie pour réduire le risque d'erreur, notamment dans la prise en charge médicale et psychologique nourrit ce fantasme. C'est justement l'idée que prône le mouvement intellectuel transhumaniste (James Hughes, Nick Bostrom, Ray Kurzweil).

Aujourd'hui l'adoption enthousiaste de la convergence NBIC, décuple la puissance de la recherche et permet des avancées spectaculaires qui rendent le rêve transhumaniste moins naïf qu'il en a l'air. Il est en tout cas bien réel dans la tête des grands géants de l'ingénierie informatique (google, Microsoft, Amazon, IBM, etc.), celui de transcender les limitations biologiques de l'espèce humaine en améliorant ses capacités physiques, intellectuelles et émotionnelles. Alors, et d'une manière concrète, pourrions-nous en finir avec la mort ? Pas au sens littéral mais au sens existentiel en tant que réalité vécue intérieurement et subjectivement. Pourrions-nous réduire le paradoxe inhérent à deux réalités antinomiques : vouloir vivre, devoir mourir ? Mieux gérer la cohabitation de deux forces opposées qui gouvernent notre psychisme : pulsions de vie *vs* pulsions de mort ?

Nous proposons, pour finir, un extrait du texte écrit par le philosophe anglais Max More de son vrai nom Max O'connor (1999) un fervent partisan du transhumanisme, qui s'intitule *Lettre à mère nature : amendements de la constitution humaine*. Ce texte expose le fondement de la philosophie transhumaniste, une vision qui peut paraître futuriste, utopiste et profondément régressive au sens psychanalytique du terme. Ce texte nous invite à penser et à repenser l'avenir de l'humain et à réinterroger le sens même de l'humanisme.

> […] Mère Nature, nous te sommes vraiment reconnaissants pour ce que tu as fait de nous. Il est clair que tu as fait du mieux que tu pouvais. Cependant, sauf ton respect, nous devons constater que sur de nombreux aspects tu as fait un travail médiocre pour ce qui est de la constitution humaine. Tu nous a fait vulnérables aux maladies et aux blessures. Tu nous as voués au vieillissement et à la mort – juste au moment où nous atteignons la sagesse. Tu as été mesquine dans l'étendue de ce qui a trait à la conscience de nos processus somatiques, cognitifs et émotionnels. Tu nous as limités en donnant les sens les plus affutés aux autres animaux. Nous n'avons la capacité de fonctionner que dans des conditions environnementales spécifiques. Tu nous as donné une mémoire limitée, un faible contrôle de nos impulsions, et des pulsions tribales et xénophobes. Et tu as oublié de nous donner le mode d'emploi de nous-mêmes !
>
> Ce que tu as fait de nous est magnifique, mais profondément imparfait […] Nous avons décidé qu'il était temps d'amender la constitution humaine […] Dans les décennies qui viennent nous allons mener une série de

changements de notre propre constitution, lancée avec les outils de la biotechnologie et guidée par une pensée créative et critique[…](More, 1999).

En guise de conclusion, il est clair que l'un des objectifs majeurs de la multiplication et le perfectionnement continu des IAs, est de transcender les limitations biologiques de l'espèce humaine en améliorant ses capacités physiques, intellectuelles et émotionnelles. Face à cela, l'Homme ne peut plus rebrousser chemin mais il doit redoubler de vigilance et réfléchir à instaurer un cadre éthique strict à l'usage des IAs dans le domaine de la santé.

**Bibliographie :**


- American Psychological Association (APA).(2023). Future of work. Advancing psychology in national artificial intelligence strategy. APA Advocacy.
- Dolto, F. (1998). Parler de la mort. Mercure de France.
- Kannampallil, T., Ajilore, O. A., Lv, N., et al. (2023). Effects of a virtual voice-based coach delivering problem-solving treatment on emotional distress and brain function: A pilot RCT in depression and anxiety. Translational Psychiatry, 13, 166. https://doi.org/10.1038/s41398-023-02462-x
- Kaywan, P., Ahmed, K., Ibaida, A., et al. (2023). Early detection of depression using a conversational AI bot : A non-clinical trial. PloSone, 18(2), e0279743. https://doi.org/10.1371/journal.pone.0279743
- Laranjo, L., Dunn, A. G., Tong, H. L., et al. (2018). Conversational agents in healthcare: A systematic

review. Journal of the American Medical Informatics Association,25(9),1248–
1258. https://doi.org/10.1093/jamia/ocy072

- More, M. (1999). Lettre à mère nature:Amendements de la constitution humaine (E.Taieb, Trad., 2018). Récupéré le 30octobre2024de http://emmanueltaieb.fr/2018/09/traduction-de-la-lettre-a-mere-nature-de-max-more/
- Nishida,T.,Nakazawa,A.,Ohmoto,Y.,&al.(2014). Conversational informatics:A data-intensive approach with emphasis on nonverbal communication. Springer.
- Raji, I.D., & Buolamwini, J. (2019). Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products. Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society.
- Weizenbaum, J. (1966). ELIZA—a computer program for the study of natural language communication between man and machine. Communications of the ACM, 9, 36 -45.
- Wolters, M. K., Kelly, F., & Kilgour, J. (2016). Designing a spoken dialogue interface to an intelligent cognitive assistant for people with dementia. Health Informatics Journal, 22(4), 854–866.

# La régulation de l'IA en matière de santé

***Wafa HARRAR MASMOUDI***

*Professeur de l'enseignement supérieur, agrégée de droit public à la Faculté des sciences juridiques, politiques et sociales de Tunis, relevant de l'Université de Carthage. Spécialisée en droit international et relations internationales (politique étrangère et diplomatie), droits de l'homme et analyse des politiques publiques et de gouvernance, professeur Wafa HARRAR MASMOUDI s'intéresse également à l'identification et à la résolution des problèmes juridiques en relation avec le domaine de la santé et de l'éthique. Ses recherches et publications portent sur des thématiques de droit international, droit de la santé, droit constitutionnel et droits de l'homme.*

*Professeur Wafa HARRAR MASMOUDI est membre du Comité national d'éthique médicale et expert-évaluateur auprès de l'Instance Nationale de l'Evaluation, de l'Assurance Qualité et de l'Accréditation. Outre une expertise développée en matière d'accréditation et audit des Systèmes Management de Qualité (SMQ)" ISO 9001, elle est experte en évaluation pédagogique et institutionnelle et varie les activités d'expertise, de formation sur plusieurs thèmes, en collaboration avec plusieurs organisations nationales et internationales.*

*Professeur Wafa HARRAR MASMOUDI est également directrice du Mastère de Recherches en Droit anglo-saxon (Common Law) à la Faculté des sciences juridiques, politiques et sociales de Tunis, membre élue du Conseil scientifique de la faculté et membre du Laboratoire de Droit des Relations Internationales des Marchés et des Négociations (DRIMAN) rattaché à la Faculté de droit et des sciences politiques de Tunis.*

Je vous remercie de votre attention aujourd'hui, alors que nous abordons un sujet d'une importance capitale, celui de la régulation de l'Intelligence Artificielle (IA) dans le domaine de la

santé. En effet, l'essor fulgurant de l'IA offre des perspectives prometteuses pour améliorer la précision des diagnostics médicaux, faciliter la découverte de nouvelles cibles thérapeutiques et permettre un traitement personnalisé tout en réduisant les coûts. Les nécessités médicales autant que les innovations technologiques expliquent le recours massif à l'Intelligence Artificielle. Son omniprésence impose au préalable d'en clarifier la signification.

Dans un sens large, l'IA peut être comprise comme la capacité d'une machine à reproduire certaines fonctions du cerveau humain, comme l'apprentissage, le raisonnement et la perception. Elle englobe un ensemble diversifié de technologies[10], parmi lesquelles le traitement du langage naturel ainsi que l'apprentissage automatique, et la reconnaissance visuelle, pour n'en nommer que quelques-unes.[11] Yann LeCun définit l'intelligence artificielle comme « *un ensemble de techniques permettant à des machines d'accomplir des tâches et de résoudre des problèmes normalement réservés aux humains et à certains animaux.* »[12]

La Recommandation sur l'éthique de l'Intelligence artificielle de l'UNESCO, adoptée à Paris le 23 novembre 2021, précise qu'il s'agit « *des systèmes technologiques capables de traiter l'information par un processus s'apparentant à un*

[10] Voir notamment RUSSEL Stuart et NORVIG Peter, Intelligence artificielle, Pearson Education France, 3ème édition, 2010, p.2.

[11] Directive (UE) 2019/770 du Parlement européen et du Conseil du 20 mai 2019 relative à certains aspects concernant les contrats de fourniture de contenus numériques et de services numériques : <https://eur-lex.europa.eu/legalcontent/FR/TXT/PDF/?uri=CELEX:32019L0770>

[12] LE CUN Yann, « Les enjeux de la Recherche en Intelligence Artificielle », Leçon inaugurale au Collège de France, 4 février 2016, Dossier de presse, 1. Qu'est-ce que l'intelligence artificielle, p.1.

*comportement intelligent, et comportant généralement des fonctions de raisonnement, d'apprentissage, de perception, d'anticipation, de planification ou de contrôle.* »[13]

Pour autant, il n'existe pas de définition unique et universelle de l'IA comme le souligne le texte de la Recommandation qui ne cherche pas - au demeurant - à en donner une puisque l'IA est appelée à évoluer en fonction des progrès technologiques. En effet, et comme le soulignent les spécialistes de l'IA, « (…) *l'absence d'une définition précise et universellement acceptée de l'IA a probablement aidé le domaine à croître, à s'épanouir et à progresser à un rythme toujours plus rapide. Les praticiens, les chercheurs et les développeurs de l'IA sont plutôt guidés par un sens approximatif de la direction à suivre et par l'impératif de "s'y mettre".* »[14]

Afin de corroborer le dissensus entourant la définition de l'intelligence artificielle (IA), les spécialistes de l'IA[15] ont compilé huit définitions technologiques qu'ils ont classées en quatre catégories : l'IA est capable de penser humainement, d'agir humainement, de penser rationnellement et d'agir rationnellement.[16] Spécialiste des nouvelles technologies, Matthew U. Scherer a - quant à lui - opté pour une définition plus pragmatique. L'IA désignerait ainsi « *les machines capables*

[13] Recommandation sur l'éthique de l'Intelligence artificielle de l'UNESCO adoptée à Parisle 23 novembre 2021 : <https://unesdoc.unesco.org/ark:/48223/pf0000380455_fre>
[14] STONE Peter et al.,Artificial Intelligence and Life in 2030, One Hundred Year Study on Artificial Intelligence: Report of the 2015–2016 Study Panel, California, Stanford University, September 2016, p.12 : "the lack of a precise, universally accepted definition of AI probably has helped the field to grow, blossom, and advanceat an ever-accelerating pace. Practitioners, researchers, and developers of AI are instead guided by a rough sense of direction and an imperative to "get on withit."".C'est nous qui traduisons.
[15] RUSSELL Stewart & NORVIG Peter, Artificial Intelligence: A Modern Approach, 3rd Ed., Prentice Hall, UpperSaddle River, NJ, 2010, pp.02-05.
[16] Idem, pp.02-05.

*d'accomplir des tâches qui, si elles étaient accomplies par un être humain, nécessiteraient de l'intelligence.* »[17]

D'aucuns ont préféré retenir l'expression « systèmes d'IA ». Il en est ainsi du Comité sur l'Intelligence Artificielle de la Chambre Haute du Parlement du Royaume-Uni qui a employé l'expression « systèmes d'IA » pour désigner les « *(…) technologies ayant la capacité d'effectuer des tâches qui nécessiteraient autrement l'intelligence humaine, telles que la perception visuelle, la reconnaissance vocale et la traduction des langues. Les systèmes d'IA ont la capacité d'apprendre ou de s'adapter à de nouvelles expériences ou à de nouveaux stimuli.* »[18] D'ailleurs, l'utilisation de l'expression « systèmes d'IA » permet de distinguer l'IA en tant que technologie tangible[19] de l'IA en tant que concept.

Ceci étant clarifié, les expressions « IA » et « systèmes d'IA » sont usitées au sens large pour englober toutes les terminologies telles que les robots, les agents autonomes, les logiciels intelligents, etc.

Les professions de santé sont la cible de prédilection des « systèmes d'IA ». Les dits systèmes sont fréquemment sollicités pour améliorer la précision du diagnostic médical[20]. En effet, ils

---

[17] SCHERER U. Matthew, "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies", in. Harvard Journal of Law & Technology, Vol. 29, No. 2, 2016, p.362 : ""Artificial intelligence" refers to machines that are capable of performing tasks that, if performed by a human, would be said to require intelligence". C'est nous qui traduisons.

[18] "AI in the UK: Ready, Willing and Able?", House of Lords Report, 2018, House of Lords Select Committee on Artificial Intelligence, Report of Session 2017-2019 : <https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>

[19] SCHERER U. Matthew, op. cit., p.362.

[20] Diagnostic assisté par IA.

permettent d'analyser les images médicales[21], et comme le souligne Diane de Saint-Affrique, « *dans certains domaines où l'analyse d'images est nécessaire, comme en radiologie diagnostique, en dermatologie ou en ophtalmologie par exemple (...), l'IA peut (...) se substituer à l'intelligence humaine et obtenir des résultats probants. Certains diagnostics pourront même être posés de façon plus rapide et plus fiable ce qui va permettre une prise en charge plus efficace des patients.* »[22] Les systèmes d'IA permettent aussi d'analyser les résultats de laboratoire, et les données recueillies dans le cadre des dossiers médicaux informatisés (DMI), en vue d'identifier des modèles permettant de prédire la présence ou la progression d'une maladie. Par ailleurs, le Règlement (UE) 2017/745 du Parlement européen et du Conseil en date du 5 avril 2017 avait auparavant reconnu le rôle de plus en plus prépondérant de l'IA dans le domaine de la recherche médicale. Celle-ci permet de prédire l'efficacité potentielle de nouvelles thérapies et aide à la conception de nouveaux médicaments.

La télémédecine constitue en outre un domaine particulièrement convoité. En effet, la télémédecine a acquis une nouvelle dimension grâce au développement de l'IA. Les chatbots médicaux, par exemple, peuvent aider à l'évaluation initiale des symptômes du patient à distance, améliorant ainsi l'accessibilité des soins dans les zones les plus reculées. Eric TOPOL a parfaitement résumé l'usage des systèmes d'IA en matière de

[21] Voir AYACHE Nicholas, « L'imagerie médicale à l'heure de l'intelligence artificielle », in. VILLANI Cédric et NORDLINGER Bernard (dir.), Santé et intelligence artificielle, CNRS Editions, oct. 2018, pp.151-154.

[22] DE SAINT-AFFRIQUE Diane, « Intelligence artificielle et médecine: quelles règles éthiques et juridiques pour une IA responsable ? », in. Médecine & Droit, 2022, vol. 2022, n° 172, pp.05-07.

santé : « *(…) D'abord, ils rendent aisé pour les praticiens l'interprétation rapide et précise des images. S'agissant ensuite des systèmes de santé, ils permettent d'améliorer le flux de travail et de réduire les erreurs médicales. Enfin, ils facilitent aux patients l'accès à l'analyse de leurs propres données contribuant ainsi à la promotion de leur santé.* »[23] L'intelligence artificielle offre ainsi un formidable potentiel pour les professionnels de la santé, qui peuvent désormais y recourir pour améliorer la qualité des soins dispensés aux patients.

Cependant, ces avancées s'accompagnent de risques intrinsèques, et soulèvent inéluctablement des questions d'ordre éthique et pratique. Il s'agit notamment des risques liés aux erreurs potentielles de diagnostic en raison des défaillances de l'IA, d'atteintes à la vie privée des patients dues à des failles de sécurité, sans pour autant occulter le risque d'exclusion de personnes n'ayant pas accès à ces technologies. En effet, « *(…) le risque, bien identifié, est que des algorithmes de décision d'apprentissage proposant des aides automatiques à la décision viennent renforcer ces biais, en ajoutent d'autres et donc discriminent.* »[24] Par ailleurs, les questions juridiques relatives à la responsabilité en cas d'erreurs médicales liées à l'IA constituent un sujet de préoccupation. Dans une lettre ouverte en date du 22 mars 2023, de nombreux chercheurs et scientifiques ont appelé à « *mettre en pause* » le développement de l'IA, en raison des

[23] TOPOL Eric J., "High-performance medicine: the convergence of human and artificial intelligence", in. Nature Medecine, n°25, 2019, pp.44-56. Abstract: " (…) for clinicians, predominantly via rapid, accurate image interpretation; for health systems, by improvingworkflow and the potential for reducing medicalerrors; and for patients, by enablingthem to process their own data to promote health". C'est nous qui traduisons.

[24] BESSE Philippe, BESSE-PATIN Aurèle, CASTETS-RENARD Céline, « Implications juridiques et éthiques des algorithmes d'intelligence artificielle dans le domaine de la santé », in. Statistique et Société, n°8, 2020, p.22.

« *profonds risques pour la société et l'humanité.* »[25] Des interrogations inédites se posent donc s'agissant de l'impact de l'intelligence artificielle sur la relation médecin-malade, et des responsabilités qui en découlent. Certains s'interrogent même sur la possibilité d'envisager la responsabilité propre de l'intelligence artificielle.

Pour autant, et pendant longtemps, l'expression n'a pas été saisie par le droit. Pourquoi ? Comme le souligne le rapport du Conseil de l'Europe[26] en date de décembre 2020, « *les définitions juridiques sont encore plus difficiles à établir que les définitions purement scientifiques. Elles doivent en effet satisfaire à plusieurs critères*[27] *(être inclusives, précises, compréhensibles, pratiques et permanentes), dont certains sont obligatoires et d'autres considérés comme des bonnes pratiques.* »[28] Certes, d'autres

[25] Pause Giant AI Experiments: An Open Letter, Future of Life Institute : <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

[26] CAHAI, Comité ad hoc sur l'intelligence artificielle (BEN-ISRAEL Isaac, CERDIO Jorge, EMA Arisa, FRIEDMAN Leehe, IENCA Marcelo, MANTELERO Alessandro, MATANIA Eviatar, MULLER Catelijne, SHIROYAMA Hideaki, VAYENA Effy), Vers une régulation des systèmes d'IA, Perspectives internationales sur l'élaboration d'un cadre juridique fondé sur les normes du Conseil de l'Europe dans le domaine des droits de l'Homme, de la démocratie et de l'Etat de droit, Titre I, Chapitre I, i. « Définition de l'IA à des fins juridiques », 2020, p.24.

[27] SCHUETT Jonas, "Defining the scope of AI Regulations", in. Law, Innovation and Technology, Volume 15 Issue 1, 2023, pp.60–82 : « Les définitions juridiques doivent être : 1) inclusives : les buts de la réglementation ne doivent pas concerner un public trop vaste ou trop limité ». ; « 2) précises : la limite entre les cas qui relèvent de la définition et ceux qui n'en relèvent pas doit être claire ; 3) compréhensibles : ceux auxquels la réglementation s'applique doivent comprendre la définition ; 4) pratiques : les professionnels du droit devraient pouvoir déterminer aisément si un cas relève ou non de la définition ; 5) permanentes : une définition juridique ne devrait pas nécessiter de constantes mises à jour ». L'auteur se réfère à l'ouvrage de BALDWIN Robert, CAVE Martin et LODGE Martin, Understanding Regulation: Theory, Strategy, and Practice, Oxford University Press, 2011, p.70.

[28] Vers une régulation des systèmes d'IA, Perspectives internationales sur l'élaboration d'un cadre juridique fondé sur les normes du Conseil de l'Europe dans le domaine des droits de l'Homme, de la démocratie et de l'Etat de droit, op. cit., p.24, note de bas de

expressions existent à l'instar de celles de « facultés mentales », de « facultés de discernement » ou encore des vocables comme la « volonté » ou le « consentement » qui renvoient nul doute à l' « intelligence », celle des sujets de droit. Cependant, c'est seulement à ce titre que celles-ci sont appréhendées par le droit positif.

Vous n'êtes pas sans savoir que le droit est fondé sur la subjectivité de l'intelligence humaine, ou en d'autres termes, sur une intelligence profondément liée à une conscience, des émotions, des sentiments, alors que l'intelligence artificielle est purement objective.[29] Cela est de nature à exacerber la difficulté à identifier la faute lorsque celle-ci est le fruit d'une intelligence purement objective.[30] Les règles juridiques créées pour l'homme sont-elles adaptées pour encadrer et réguler les risques liés aux systèmes d'IA sans pour autant « brider » l'innovation ?Il s'agira de trouver « *un point d'équilibre entre flexibilité – pour accompagner les progrès techniques – et précision – pour apporter la sécurité juridique nécessaire.* »[31] C'est le choix cornélien auquel font face les pouvoirs publics de par le monde.[32]

---

page n°25 : « L'obligation d'inclusivité découle du principe de proportionnalité, en droit de l'UE (article 5.4 du traité sur l'Union européenne). Toujours en droit de l'UE, les critères de précision et de caractère compréhensible reposent sur le principe de sécurité juridique. La praticité et la permanence sont considérées comme de bonnes pratiques législatives. »

[29] Voir le projet de recherche initié par John Mc CARTHY et Marvin MINSKY en 1956 dont l'objectif est de reproduire les facultés cognitives de l'intelligence humaine au moyen de l'ordinateur.

[30] Déceler la « volonté » en droit civil implique de révéler l'élément « conscience », de la même manière que l'élément « intention » pour ce qui est du droit pénal.

[31] MAZEAU Laurène, « Responsabilité. Quel(s) droit(s) pour les systèmes d'intelligence artificielle ? », in. Cahiers Droit, Sciences & Technologies, 2020, n° 10, pp. 223-229.

[32] Force est d'admettre que« les autorités sanitaires et le secteur public sont (donc) devant un choix cornélien : réguler fortement l'utilisation de l'IA en santé et prendre un

Le cadre législatif et réglementaire relatif à l'utilisation de l'intelligence artificielle n'en est certes qu'à ses balbutiements. Il n'en demeure pas moins que certaines initiatives commencent à porter leur fruit, et ce, à l'instar du récent décret relatif au développement et utilisation sécurisée et fiable de l'intelligence artificielle[33] édicté par le président Joe BIDEN en 2023. Ledit décret établit des normes inédites aux États-Unis en matière de sécurité, de protection de la vie privée et de droits civils, tout en mettant l'accent sur la nécessité d'encourager et de stimuler l'innovation et le progrès scientifiques.

A l'échelle du continent européen, les trois institutions de l'Union européenne, à savoir le Parlement, le Conseil et la Commission, sont parvenues à un accord historique portant sur l'IA le 8 décembre 2023.[34] Le texte a été officiellement adopté par le Conseil de l'Union européenne le 21 mai 2024 avant d'être signé le 13 juin 2024 par le président du Parlement européen et le président du Conseil de l'Union européenne. Ce texte - le premier à encadrer juridiquement l'intelligence artificielle - constitue la première règlementation contraignante de l'intelligence artificielle.

---

retard considérable ou ouvrir grand les portes du marché du soin aux solutions d'intelligence artificielle ce qui risque de favoriser l'émergence de système défectueux voire dangereux ce qui posera des problèmes important en terme de responsabilité et d'éthique ». Professeur BIBAULT Jean-Emmanuel, « Comment réguler l'intelligence artificielle en médecine ? », in. Journal Le Figaro en date du 2 mai 2019 : <https://sante.lefigaro.fr/article/comment-reguler-l-intelligence-artificielle-en-medecine>

[33] The White House. Executive order on the safe, secure, and trust worthy development and use of AI. 2023 : <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>

[34] Proposition de Règlement du Parlement européen et du Conseil établissant des Règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union européenne : <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX%3A52021PC0206>

La loi sur l'IA a été publiée au Journal officiel de l'Union européenne le 12 juillet 2024, et est entrée en vigueur le 1er août 2024. La date d'entrée en vigueur de la loi relative à l'intelligence artificielle marque le début de son application. À l'heure où nous nous exprimons, aucune des exigences posées par cette législation n'est mise en œuvre ; leur application interviendra de manière progressive conformément à un calendrier prédéfini.[35] C'est ainsi que « *l'UE devient le premier continent à fixer des règles claires pour l'utilisation de l'IA*. »[36] Aujourd'hui le 2 novembre 2024 constitue le terme imparti aux États membres pour identifier et publier la liste des autorités et organismes chargés de la protection des droits fondamentaux. Les États membres sont tenus, en outre, d'informer la Commission ainsi que les autres États membres de cette démarche.

Le Règlement relatif à l'IA adopte une approche innovante puisque ce n'est pas seulement la technologie *per se* qui est encadrée, mais ses cas d'usage. Le système d'intelligence artificielle (système d'IA) y est défini comme « *un système automatisé qui est conçu pour fonctionner à différents niveaux d'autonomie et peut faire preuve d'une capacité d'adaptation après son déploiement, et qui, pour des objectifs explicites ou*

[35] Article 113 : « Le présent règlement entre en vigueur le vingtième jour suivant celui de sa publication au Journal officiel de l'Union européenne. Il s'applique à partir du 2 août 2026.
Cependant :
(a) Les chapitres I et II s'appliquent à partir du 2 février 2025 ;
(b) Le chapitre III, section 4, le chapitre V, les chapitres VII et XII et l'article 78 sont applicables à partir du 2 août 2025, à l'exception de l'article 101.
(c) L'article 6, paragraphe 1, et les obligations correspondantes du présent règlement s'appliquent à partir du 2 août 2027. ».

[36] BRETON Thierry, Commissaire européen à l'origine du projet qui a été présenté au mois d'avril 2021, « AI Act : l'Union européenne trouve « un accord politique » », publié le 11 Décembre 2023 (Synthèses de presse).

*implicites, déduit, à partir des entrées qu'il reçoit, la manière de générer des sorties telles que des prédictions, du contenu, des recommandations ou des décisions qui peuvent influencer les environnements physiques ou virtuels.* »[37]

Partant du postulat que le système d'IA constitue en soi une entité juridique indéterminée, son usage en médecine pose plusieurs difficultés. Si l'on considère qu'il y a une grande similarité entre les capacités cognitives humaines et les systèmes d'intelligence artificielle, voire une équivalence, comme le soutiennent certains, il conviendra d'admettre que les règles juridiques « *pensées en considération de l'intelligence humaine* » puissent être appliquées *mutatis mutandis* à son homologue, IA. Au contraire, si l'on considère qu'il existe « *une différence de nature entre l'homme et la machine* », il conviendra à ce moment-là d'inventer de nouvelles règles.

Réfléchir à un régime juridique applicable à l'IA en matière de santé n'est pas chose aisée. Il s'agira de s'interroger dans quelle mesure le droit positif permet-il d'assurer un encadrement efficace de l'utilisation des systèmes d'intelligence artificielle en matière de santé, tant au regard des obligations préventives qui en gouvernent l'usage qu'au regard des régimes de responsabilité appelés à se déployer en cas de défaillance ?

Pour ce faire, il conviendra d'examiner, en premier, les obligations découlant de l'usage des systèmes d'IA en matière de santé, celles-ci sont modulées en fonction des risques (I) avant de tenter d'identifier les règles applicables en matière de

[37] Article 3 du Règlement du Parlement européen et du Conseil établissant des Règles harmonisées concernant l'intelligence artificielle (législation relative à l'intelligence artificielle) et modifiant certains actes législatifs de l'Union européenne : <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CONSIL:PE_24_2024_INIT&qid=1717966367678>

responsabilité, une responsabilité qui demeure difficile à circonscrire (II).

## I - Des obligations modulées en fonction des risques

Les obligations des opérateurs[38] des systèmes d'IA varient en fonction des risques encourus par les patients. Le cadre juridique semble se préciser grâce aux récentes initiatives prises au niveau européen, ce qui permet désormais de procéder à une identification avertie des risques, et d'établir des corrélations pour en tirer les obligations qui en découlent.[39] En effet, l'évolution rapide des systèmes d'intelligence artificielle a conduit à une prise de conscience accrue des risques inhérents à leur déploiement. Ces risques, qui varient en fonction de la nature, des applications et des contextes d'utilisation des systèmes d'IA, ont imposé une approche spécifique en matière de régulation.

L'Union européenne a élaboré un cadre normatif instaurant un régime d'identification des risques visant à classer les SIA selon leur potentiel de dangerosité. Il s'agit d'une approche graduée fondée sur l'identification des risques qui s'articule ainsi autour d'une échelle de niveaux de dangerosité.

---

[38] Règlement (UE) 2024/1689 relatif à l'Intelligence artificielle : Règlement du Parlement européen et du Conseil établissant des Règles harmonisées concernant l'intelligence artificielle (législation sur l'intelligence artificielle) et modifiant certains actes législatifs de l'Union européenne, tel qu'adopté par le Conseil de l'Union européenne le 21 mai 2024, signé le 13 juin 2024 par le président du Parlement européen et le président du Conseil de l'Union européenne, Article 3, point (8) « opérateur », le fournisseur, l'utilisateur, le mandataire, l'importateur et le distributeur.

[39] « La classification d'un système d'IA comme étant à haut risque repose sur la finalité du système d'IA, conformément à la législation existante en matière de sécurité des produits. Par conséquent, la classification d'un système d'IA comme étant à haut risque ne dépend pas seulement de la fonction remplie par le système d'IA, mais également de la finalité et des modalités spécifiques pour lesquelles ce système est utilisé ». Loi sur l'Intelligence artificielle (Règlement (UE) 2024/1689, version du Journal officiel de l'UE du 13 juin 2024, entrée en vigueur le 1er août 2024.

Cette démarche s'inscrit dans le cadre de la réglementation prudentielle, qui a déjà été appliquée notamment en matière de protection des données personnelles.

Ce cadre normatif distingue quatre niveaux de risques allant des SIA à risque inacceptable, les SIA à haut risque, les SIA à risque limité aux SIA à risque minimal. Cette classification s'appuie sur des critères objectifs, notamment l'impact potentiel sur les droits fondamentaux, la sécurité et la santé des personnes.[40]

Le risque qualifié d'inacceptable correspond, de manière générale, au risque dont le niveau de dangerosité rend l'usage du système d'intelligence artificielle absolument interdit. Il s'inscrit dans une classe de dangers dont la portée et les répercussions potentielles sont telles qu'elles rendent leur gestion et leur utilisation radicalement incompatibles avec les principes cardinaux de la dignité humaine et de la sauvegarde des droits individuels. En effet, dans certains contextes, l'emploi d'un système d'intelligence artificielle pourrait engendrer un dommage substantiel, immédiat et flagrant envers une collectivité d'individus, en tirant parti de leurs fragilités d'une manière qui serait considérée comme éthiquement condamnable et juridiquement inacceptable.

A dessein, le risque qualifié d'inacceptable renvoie à des situations où l'exploitation d'un système IA pourrait exploiter la vulnérabilité d'un groupe de personnes dans le but de leur causer un dommage manifeste. La législation européenne a ainsi identifié certains systèmes d'intelligence artificielle, qualifiés de risques

[40] Règlement (UE) 2024/1689 relatif à l'Intelligence artificielle.

inacceptables, dont l'utilisation est prohibée.[41] Parmi ces systèmes, figurent la manipulation cognitive du comportement, la police prédictive, la reconnaissance des émotions dans les environnements de travail ou au sein des établissements scolaires, ainsi que l'évaluation sociale. L'identification du caractère « inacceptable » du risque est par conséquent intrinsèquement liée à la nature des préjudices qui en découlent. Ces préjudices se manifestent généralement par des atteintes graves à la liberté individuelle, à la vie privée ou encore à l'intégrité psychologique et sociale des individus ou des groupes concernés. Ce type de risque implique, par définition, une exploitation technologique dont les effets nuisibles sont d'une telle ampleur qu'ils rendent toute forme d'exploitation du système d'intelligence artificielle dans ces contextes interdite. Ces interdictions visent à préserver les individus, tant sur le plan moral que sur le plan juridique.

L'approche graduée fait ensuite état des systèmes d'intelligence artificielle présentant un risque élevé (SIA à haut risque). Un système d'intelligence artificielle à haut risque se définit comme tout dispositif technologique autonome dont l'usage constitue un danger potentiel significatif pour les individus, un danger dont les conséquences sont imprévisibles et susceptibles de se manifester de manière contraire aux attentes raisonnablement fondées sur les usages ordinaires de la technologie. Le critère fondamental de cette qualification repose sur le degré de gravité des préjudices que ledit système pourrait causer, qu'il s'agisse de préjudices affectant la santé ou la

[41] Chapitre II, Article 5 du Règlement du Parlement européen et du Conseil établissant des Règles harmonisées concernant l'intelligence artificielle (législation relative à l'Intelligence artificielle) et modifiant certains actes législatifs de l'Union européenne : <https://eur-lex.europa.eu/legal-content/FR/TXT/PDF/?uri=CONSIL:PE_24_2024_INIT&qid=1717966367678>.

sécurité.[42] Les caractéristiques d'un tel système sont multiples et reposent principalement sur trois éléments déterminants. Premièrement, l'autonomie du système d'IA, qui lui permet de fonctionner de manière indépendante, sans intervention humaine immédiate, constitue un facteur clé d'imprévisibilité. Cette autonomie, bien que source de progrès technologique, est également porteuse de risques élevés, car elle empêche un contrôle direct et constant des décisions prises par l'IA. Deuxièmement, un risque significatif de préjudice se déploie lorsque l'utilisation du système d'IA, par ses effets directs ou indirects, peut entraîner des atteintes graves à l'intégrité physique, psychologique ou sociale des individus. Enfin, la dimension imprévisible des conséquences des actions ou décisions prises par le système d'IA représente un facteur aggravant.

Par conséquent, un système d'intelligence artificielle est considéré à haut risque notamment lorsque son utilisation autonome présente une probabilité significative de porter préjudice à des personnes, de manière imprévisible et en-dehors des normes raisonnablement attendues.[43] À cet effet, le Chapitre III[44] de la législation européenne comporte des

[42] Synthèse à partir d'une lecture croisée des dispositions de l'article 6 ainsi que de l'Annexe III du Règlement relatif à l'Intelligence artificielle.

[43]Une liste exhaustive de tous les systèmes d'IA à haut risque est annexée au Règlement. La qualification du caractère des risques (élevés, très élevés) figure dans une liste annexée au Règlement.

[44] Article 6 : Règles relatives à la classification de systèmes d'IA comme systèmes à haut risque :

« 1. Un système d'IA mis sur le marché ou mis en service, qu'il soit ou non indépendant des produits visés aux points a) et b), est considéré comme à haut risque lorsque les deux conditions suivantes sont remplies:

(a) le système d'IA est destiné à être utilisé comme composant de sécurité d'un produit couvert par les actes législatifs d'harmonisation de l'Union énumérés à l'annexe II, ou constitue lui-même un tel produit;

dispositions particulières applicables en matière de santé, renvoyant de surcroît à l'Annexe III, notamment au point 5, qui traite de l' « Accès et droit aux services privés essentiels et aux services publics et prestations sociales essentiels »[45]. Il ressort en outre de l'article 7 (b) que les systèmes d'IA sont considérés à haut risque lorsqu'ils présentent un risque tangible pour la santé et la sécurité des individus, ou lorsqu'ils sont susceptibles d'avoir des conséquences négatives sur les droits fondamentaux des personnes[46], notamment la vie privée[47], l'égalité et la non-discrimination.

A côté des systèmes d'intelligence artificielle (SIA) présentant des risques qualifiés d'inacceptables, ainsi que des SIA à haut risque, le cadre normatif établit un troisième niveau, en

---

(b) le produit dont le composant de sécurité est le système d'IA, ou le système d'IA lui-même en tant que produit, est soumis à une évaluation de la conformité par un tiers en vue de la mise sur le marché ou de la mise en service de ce produit conformément aux actes législatifs d'harmonisation de l'Union énumérés à l'annexe II.

2. Outre les systèmes d'IA à haut risque visés au paragraphe 1, les systèmes d'IA visés à l'annexe III sont également considérés comme à haut risque. ».

[45] Annexe III (5) « Accès et droit aux services privés essentiels et aux services publics et prestations sociales essentiels:

a) systèmes d'IA destinés à être utilisés par les autorités publiques ou en leur nom pour évaluer l'éligibilité des personnes physiques aux prestations et services d'aide sociale essentiels, y compris les services de soins de santé, ainsi que pour octroyer, réduire, révoquer ou récupérer ces prestations et services ; » ainsi que « c) systèmes d'IA destinés à être utilisés pour l'évaluation des risques et la tarification en ce qui concerne les personnes physiques en matière d'assurance-vie et d'assurance maladie ; d) systèmes d'IA destinés à évaluer et hiérarchiser les appels d'urgence émanant de personnes physiques ou à être utilisés pour envoyer ou établir des priorités dans l'envoi des services d'intervention d'urgence, y compris par la police, les pompiers et l'assistance médicale, ainsi que pour les systèmes de tri des patients admis dans les services de santé d'urgence ».

[46] Conformément à l'Article 7 (b), il s'agit des systèmes d'IA qui « présentent un risque de préjudice pour la santé et la sécurité, ou un risque d'incidence négative sur les droits fondamentaux ».

[47] Voir DONNAT Francis, « L'intelligence artificielle, un danger pour la vie privée ? », in. Pouvoirs, n°170, septembre 2019, Seuil, pp.95-103.

l'occurrence, les SIA présentant un risque limité. Celui-ci se réfère aux dangers liés au manque de transparence qui pourraient résulter de l'utilisation des systèmes d'intelligence artificielle.

Enfin, il existe un dernier niveau de risques, en l'occurrence, le risque minimal, où aucune règle spécifique n'est exigée. Dans ce cadre, le Règlement relatif à l'intelligence artificielle permet l'utilisation libre de systèmes d'IA présentant des risques minimaux. Il peut s'agir par exemple des applications qui incluent des outils légers comme les jeux vidéo ou les filtres anti-spam, et force nous est de reconnaître que - pour l'heure - plusieurs systèmes d'intelligence artificielle employés peuvent être rangés parmi ce niveau de risques.

L'objectif de cette approche nuancée adoptée par la législation européenne est d'établir un cadre de précaution, visant à limiter les risques inhérents aux systèmes d'IA. Mais l'identification des différents niveaux de risque des systèmes d'IA ne se contente pas d'offrir un intérêt strictement didactique, puisqu'elle a des effets directs sur les obligations des différents acteurs, qu'ils soient fournisseurs, importateurs, distributeurs, ou encore déployeurs de systèmes d'IA. En effet, ces obligations sont modulées en fonction des risques identifiés ; elles doivent, en d'autres termes, être comprises dans l'échelle de risques décrites ci-haut.

Le Règlement relatif à l'IA introduit une obligation de transparence vis-à-vis des utilisateurs. La transparence désigne le degré d'informations fournies à l'utilisateur, incluant la structure du modèle, ses usages prévus, les décisions de déploiement (comment, quand et par qui), ainsi que les choix de conception

et les données d'entraînement.[48] Appliquée au domaine de la santé, l'obligation de transparence peut être considérée comme un gage de confiance des patients envers le système d'IA qui doit leur être expliqué de manière à ce qu'ils puissent accéder à des informations détaillées et ainsi comprendre le diagnostic ou la proposition de traitement.

L'obligation de transparence pèse en premier sur le concepteur du dispositif IA qui doit assurer une « explicabilité » aux utilisateurs. Dans le domaine de l'Intelligence Artificielle, « *l'explicabilité est la capacité de mettre en relation et de rendre compréhensible les éléments pris en compte par le système d'IA pour la production d'un résultat. Les explications doivent être adaptées au niveau de compréhension de la personne à laquelle elles sont destinées.* »[49]

Afin d'optimiser le déploiement de l'IA dans les environnements de soins de santé, l'obligation de transparence pèse également sur les développeurs de systèmes d'IA en vue d'instaurer une relation synergique entre les professionnels de la santé et les experts en technologie. Cette collaboration est essentielle pour garantir que les recommandations formulées par l'IA soient méticuleusement examinées pour en vérifier l'exactitude, en vue de minimiser le risque d'erreurs pouvant découler de données erronées ou d'algorithmes biaisés. Une conférence-réaction en date du 22 Janvier à Paris a porté sur les

[48] Voir notamment MALEK Md. Abdul, « Transparency in Predictive Algorithms : A Judicial Perspective », 2021, disponible sur : <https://www.researchgate.net/publication/352145154_Transparency_in_Predictive_Algorithms_A_Judicial_Perspective>

[49] Commission Nationale de l'Informatique et des Libertés, Consultable sur : <https://www.cnil.fr/fr/definition/explicabilite-ia#:~:text=Dans%20le%20domaine%20de%20l,la%20production%20d'un%20r%C3%A9sultat>

algorithmes capables de diagnostiquer ou prédire les évolutions d'une maladie[50] grâce aux progrès récents de reconnaissance d'image permettant de repérer efficacement des anomalies dans le corps. Lors de cette conférence, la question critique de la relation entre médecine et système d'IA/algorithme s'est posée notamment avec acuité en contexte de pandémie s'agissant des algorithmes cliniques utilisés pour l'identification des personnes à risque de vulnérabilité psychosociale.[51]

Outre l'obligation de transparence, une obligation d'information pèse sur les professionnels de santé en raison de la capacité des systèmes d'IA « *à exploiter les échanges d'informations sur la santé pour récupérer les dossiers médicaux des patients, les analyser et formuler des questions pertinentes en fonction des antécédents médicaux du patient*. »[52] Certes, le cadre juridique de l'exploitation des données de santé en Tunisie est clairement établi par la Loi organique n°2004-63 relative à la protection des données à caractère personnel[53], mais il s'agit ici

[50] Conférence-réaction en date du 22 janvier « Médecine, algorithme et responsabilité », ESPCI, Paris, PSL Dauphine, <https://www.espgg.org/Medecine-algorithme-et>

[51] BELAND Jean-Pierre, « L'intelligence artificielle : un levier permettant de restaurer l'égalité ? », in. Canadian Journal of Bioethics, Revue canadienne de bioéthique, 5(2), 2022, pp.162-170.

[52] Par exemple, en recoupant la liste des médicaments d'un patient et ses problèmes de santé actuels, l'IA peut vérifier si les patients respectent les régimes qui leur ont été prescrits ou s'ils ont cessé de prendre des médicaments contradictoires à la lumière de nouvelles données médicales. Ce processus permet d'établir un historique médical plus complet pour le patient, qui peut ensuite être utilisé par les médecins pour fournir de meilleurs soins.

[53] Loi organique n°2004-63 du 27 juillet 2004 portant sur la protection des données à caractère personnel et le Décret n° 2007-3004 du 27 novembre 2007, fixant les conditions et les procédures de déclaration et d'autorisation pour le traitement des données à caractère personnel. Le traitement des données à caractère personnel est interdit (Article 14 de la loi-organique de 2004). L'article 14 commence en effet par interdire le traitement des données personnelles concernant l'origine génétique de la personne ainsi que celles relatives à sa santé. Le traitement est autorisé lorsqu'il est effectué avec le consentement exprès de la personne concernée donné par n'importe que

d'imposer une véritable obligation de transparence des algorithmes utilisés par l'intelligence artificielle, afin de conserver la capacité de contrôler cette dernière, tant pour l'opérateur que pour le patient (potentiel) concerné. Il convient de rappeler que l'obligation de transparence pesant sur les algorithmes est d'ores et déjà consacrée dans un certain nombre de pays, même si la question de son effectivité n'est pas encore garantie. En Tunisie, le projet Code du numérique approuvé par le Conseil des

---

moyen laissant une trace écrite, ou lorsque ces données ont acquis un aspect manifestement public, ou lorsque de traitement s'avère nécessaire à des fins historiques ou scientifiques ou, lorsque ce traitement est nécessaire à la sauvegarde des intérêts vitaux de la personne concernée. La loi-organique n°2004-63 du 27 juillet 2004 s'applique au traitement automatisé ainsi qu'au traitement non automatisé des données de santé à caractère personnel mis en œuvre par des personnes physiques ou par des personnes morales (Chapitre V - Section II – Traitement des données à caractère personnel relatives à la santé). Il est important de relever que le traitement d'une donnée de santé nécessite dans le cas courant une simple déclaration, lorsqu'elle doit être diffusée au public, il devient nécessaire - exception - d'obtenir le consentement de la personne concernée mais – exception de l'exception - ce consentement devient inutile en cas d'urgence.

De manière plus précise, il conviendra de considérer les conditions suivantes posées par la loi-organique de 2004 : La personne physique ou le représentant légal de la personne morale désirant effectuer le traitement des données à caractère personnel ainsi que leurs agents doivent remplir les conditions suivantes : ils doivent être de nationalité tunisienne et résider en Tunisie ; ils ne doivent pas avoir de casier judiciaire. Il est impératif de noter que ces conditions s'appliquent également aux sous-traitants et à leurs agents.

Le traitement des données à caractère personnel en matière de santé est réalisé par les établissements publics de santé (l'article 53, 1erparag. se réfère également aux autorités publiques, collectivités locales et établissements publics à caractère administratif). L'on remarquera ainsi qu'une donnée de santé traitée par une personne publique n'a besoin ni d'autorisation ni même de déclaration, ni du consentement préalable de la personne concernée. Quand un traitement est bénéfique à la santé de la personne ou lorsqu'il concerne des données génétiques, un régime juridique différent, plus souple, est établi. Ceci étant clarifié, le traitement des données à caractère personnel relatives à la santé ne peut être mis en œuvre que par des médecins ou des personnes soumises, en raison de leur fonction, à l'obligation de garder le secret professionnel. Dans le cadre des recherches scientifiques en matière de santé, les professionnels de santé sont légalement autorisés à divulguer des données à caractère personnel à des personnes ou des institutions engagées dans ce type de recherche, à la demande des entités susmentionnées et sous réserve d'une autorisation de l'Instance Nationale de Protection des Données à Caractère Personnel.

ministres et introduit au Parlement depuis le 26 décembre 2019 n°80/2019 n'a pas encore vu le jour.

Au niveau de l'Union européenne, la Section 2 du Règlement relatif à l'Intelligence artificielle, et notamment l'article 13[54] impose des obligations de transparence et information à certains systèmes d'IA en raison des risques spécifiques de manipulation qu'ils présentent. C'est à ce titre que les systèmes d'IA, notamment en matière de santé, sont concernés. En effet, l'obligation de transparence devra s'appliquer aux systèmes qui, de manière cumulative ou alternative, interagissent directement avec des personnes physiques, analysent des émotions ou établissent des corrélations avec des catégories sociales à partir de données biométriques, ou encore génèrent ou modifient des contenus. Lorsque des personnes interagissent avec un système d'IA ou que leurs émotions ou caractéristiques sont reconnues par des moyens automatisés, elles doivent en être non seulement informées mais elles doivent également donner leur consentement.[55]

Le principe du consentement éclairé, essentiel dans la pratique médicale, est mis à l'épreuve par l'usage des systèmes d'IA, ce qui serait susceptible de rendre la prise de décision médicale plus complexe. Les patients doivent être pleinement informés du rôle que joue le système d'IA dans leur parcours de santé et de ses implications en matière de protection de la vie

[54] Règlement relatif à l'Intelligence artificielle (Règlement (UE) 2024/1689, version du Journal officiel de l'UE du 13 juin 2024. L'article entrera en vigueur le 2 août 2026 (conformément aux dispositions de l'article 113).

[55] STANKOVICH Miriam, "Regulating AI and Big Data Deployment in Healthcare: Proposing Robust and Sustainable Solutions for Developing Countries' Governments", in. Global Toolkit on AI and the Rule of Law for the Judiciary, UNESCO, 2023, disponible sur : <https://unesdoc.unesco.org/ark:/48223/pf0000387331>

privée en particulier lorsque la collecte de données va au-delà des dossiers médicaux traditionnels pour inclure des informations provenant d'autres sources, de dispositifs portables et de technologies intelligentes. Il est donc clair que l'obligation d'information s'applique ainsi aux professionnels de la santé. En la matière, ce sont les règles de droit commun qui demeurent applicables, et ce, en l'absence d'un régime juridique particulier aux systèmes d'IA. Le manquement à l'obligation d'information est assimilé à une faute[56], susceptible d'être qualifié de « faute d'humanisme », par opposition à la faute technique. Il s'agit d'une faute commise pendant l'accomplissement de l'acte médical, qu'elle survienne à l'occasion du diagnostic ou lors de la détermination des soins et thérapies nécessaires.

La jurisprudence est départagée s'agissant de la question de réparation : la jurisprudence tunisienne considère que cette faute nécessite une réparation de tout le dommage corporel subi[57] tandis que la jurisprudence française indemnise au titre d'une réparation pour « perte de chance ».[58]Dans ce cas de figure, les juges indemnisent ce qu'ils considèrent être une perte de chance d'éviter l'opération due au manquement à l'obligation d'information.[59] Le défaut d'information serait ainsi à l'origine

[56] Voir notamment VIALLA François, « La défaillance dans l'information : une faute d'humanisme », in. Revue juridique de l'Océan Indien, 2013, pp.09-29.

[57] Arrêts des juges de fond C.A. de Tunis n°48788 du 29 avril 1998, in. Revue Tunisienne de Droit, 1999, Jurisprudence, 1) Droit civil, b. Obligations et Contrats, p.201 et suiv., C.A de Tunis n°95747 du 4 juin 2003, in. Revue de la Jurisprudence et de la Législation, janvier 2005, p.137. Arrêt de la Cour de cassation civile n°20241.2007, du 13 mars 2008 (inédit).

[58] GUETTIER Christophe, « Perte de chance de survie ou de guérison », in. *Revue juridique de l'Ouest*, 2007-4. pp. 311-329.

[59] [Jurisprudence] Obligation d'information et perte de chance (à propos du durcissement de la responsabilité médicale), in. La lettre juridique, décembre 2005, consultable sur :

d'une perte de chance d'éviter l'acte dommageable.[60] L'utilisation par les professionnels de la santé des systèmes d'IA conduira sans doute à l'extension du contenu de l'obligation d'information. En effet, l'obligation d'information devra désormais porter sur « *l'utilisation qui sera faite des données personnelles du patient, voire sur les risques liés à l'utilisation de l'intelligence artificielle* »[61] au risque d'aboutir à un *abus de l'obligation d'information*. Et comme l'écrit Grégoire LOISEAU : « *En l'absence de preuve d'un défaut du système d'intelligence artificielle, la responsabilité pourrait être recherchée sur le fondement de l'obligation d'information. Ne pouvant accabler la machine, on accablerait l'homme qui n'a pas prévenu le patient de tous les risques, même exceptionnels, que l'utilisation de celle-ci pouvait faire encourir. La stimulation de l'obligation d'information pourrait alors conduire, possiblement, à en faire évoluer l'étendue.* »[62]

Nous avons tantôt évoqué la télémédecine en tant que domaine de prédilection des systèmes d'IA. La télémédecine basée sur l'IA, à l'instar des chatbots médicaux, doit également

---

<https://www.lexbase.fr/article-juridique/3207792-jurisprudence-obligation-dinformation-et-perte-de-chance-a-propos-du-durcissement-de-la-responsabili>

[60] La jurisprudence administrative tunisienne a souscrit à cette interprétation dans un arrêt de principe du TA en date du 31 décembre 2007 (TA, 3ème Chambre de cassation 31 décembre 2007, n°38643, Chef du contentieux de l'Etat c/Zohra BAZZEZI), position réitérée dans l'arrêt du TA, Appel n°27458 du 29 juin 2011. Voir notamment la communication de Dr ZINELABIDINE Taha, « Information du malade. Etat des lieux : Aspects juridique et déontologique », in. La Communication dans le domaine de la santé, XXIème Conférence annuelle du Comité national d'éthique médicale, Beït Al-Hekma, Carthage, 8 décembre 2017, p.47.

[61] LEQUILLERIER Clémentine, « L'impact de l'intelligence artificielle sur la santé », in. Journal de Droit de la santé et de l'assurance maladie, Dossier thématique : Intelligence artificielle et santé, N°17, 2017, p.19.

[62] LOISEAU Grégoire, « Intelligence artificielle et santé », Dossier thématique, in. Journal de Droit de la santé et de l'assurance maladie, n°17, 2017, p.22.

être réglementée afin de tenir compte des risques liés à l'usage des systèmes d'IA en termes de respect de la confidentialité des patients et de garantie de la sécurité et de l'efficacité des soins. Les textes organisant la télémédecine, et plus précisément l'article 23 bis de la loi n°2018-43 du 11 juillet 2018, complétant la loi n°91-21 du 13 mars 1991, relative à l'exercice et à l'organisation de la profession de médecin et de médecin-dentiste dispose que : « *le médecin ou le médecin dentiste traitant ne doit procéder à aucun acte dans le cadre de la télémédecine qu'après en avoir informé le patient, et, le cas échéant, son tuteur légal et avoir recueilli son consentement éclairé, et ce, par tout moyen laissant une trace écrite ou électronique.* »[63] Il en est de même du décret présidentiel n°318/2022 du 08 avril 2022 fixant les conditions générales d'exercice de la télémédecine et ses champs d'application[64] qui dispose que : « *avant la réalisation de tout acte de télémédecine, le consentement libre et éclairé du patient ou, le cas échéant, de son tuteur légal, doit être recueilli et ce, après son information de la nécessité, de l'intérêt, des conséquences et de la portée dudit acte ainsi que des moyens mis en œuvre pour sa réalisation* ».

Après avoir examiné - dans le cadre du premier axe de cette intervention - de manière succincte les obligations drainées par les risques inhérents à l'usage des systèmes d'IA, il convient de s'arrêter, dans le cadre du second axe, sur l'épineuse question de la responsabilité que lesdits risques risquent de poser.

---

[63] Loi n° 2018-43 du 11 juillet 2018, complétant la loi n° 91-21 du 13 mars 1991, relative à l'exercice et à l'organisation de la profession de médecin et de médecin dentiste, J.O.R.T n° 057 en date du 17 juillet 2018, p.2422.

[64] Décret Présidentiel n° 2022-318 du 8 avril 2022, fixant les conditions générales d'exercice de la télémédecine et les domaines de son application, J.O.R.T n° 040 en date du 12 avril 2022, pp.1057-1064.

## II - Une responsabilité difficile à circonscrire

Nous sommes au cœur d'une problématique qui s'annonce complexe, en l'occurrence, celle de la détermination de la responsabilité des SIA en matière de santé. En effet, la pluralité et la spécificité des systèmes d'IA imposent une réflexion sur la notion de responsabilité afin de vérifier si les règles du droit de la responsabilité civile en vigueur peuvent être appliquées lorsque l'on constate un dommage causé par le fait d'un système d'intelligence artificielle.

Il convient d'abord de sonder les fondements pour agir en cas de préjudice subi du fait d'un système d'IA. Certes, plusieurs textes existent régissant la protection de la vie privée, et les données à caractère personnel y compris les données de santé. Il existe, par ailleurs, un véritable arsenal juridique inhérent aux droits de propriété intellectuelle applicable à chaque fois que l'exploitation des données par les systèmes d'IA et les résultats qu'ils produisent contreviennent aux droits de propriété intellectuelle. Ces textes demeurent cependant insuffisants pour couvrir la question de la réparation du dommage subi du fait d'un système d'IA. Par conséquent, deux options ont été avancées. La première milite en faveur de la création d'une personnalité juridique propre aux SIA tandis que la seconde défend le recours aux règles classiques de la responsabilité moyennant des aménagements. Le choix n'est pas aisé : s'agira t-il de créer un nouveau régime spécifique dédié ou d'adapter les règles applicables en matière de responsabilité ?

La première option, bien que théoriquement séduisante, est aujourd'hui minoritaire. *De facto*, c'est la seconde qui a été retenue, notamment par la législation européenne dans le cadre du Règlement relatif à l'Intelligence artificielle. Celui-ci ne

manquera pas, nul doute, d'influencer les législations de nombreux États qui sont encore dans une phase de tâtonnements. La législation européenne souligne que la responsabilité doit demeurer ancrée dans les principes traditionnels, manifestant ainsi une tendance à la prudence face à l'innovation technologique.

Il convient néanmoins - dans le cadre de cette intervention - de s'attarder brièvement sur la première option, celle de l'attribution d'une personnalité juridique à l'IA, quand bien même celle-ci est pour l'heure écartée. En effet, nombreuses sont les voix qui se sont élevées pour revendiquer l'attribution d'une personnalité juridique propre à l'IA, par analogie avec la création - par le législateur - d'une personnalité spécifique aux personnes morales.[65] Il s'agira de reconnaître une personnalité juridique spécifique, sinon aux systèmes d'intelligence artificielle en général, du moins aux robots intelligents[66] tout en rappelant que dans la littérature, le terme robot déborde souvent sa définition sémantique pour intégrer les réflexions sur l'intelligence artificielle[67].La résolution adoptée par le Parlement européen le 16 février 2017, formulant des recommandations à la Commission concernant des règles de droit civil sur la robotique, s'en est fait l'écho en proposant une hypothétique « *personnalité électronique*. »[68] Il serait question de reconnaître une personnalité

[65] La personnalité morale est une fiction juridique portant création d'une nouvelle catégorie de personnes juridiques.

[66] Voir l'analyse de BENSOUSSAN Alain, « Droit des robots : science-fiction ou anticipation ? », in. D., 2015, p.1640.

[67] ROUVIÈRE Frédéric, « Le raisonnement par algorithmes : le fantasme du juge-robot », in. RTDCiv., Revue trimestrielle de droit civil, Dalloz, 2018, p.530 et suiv.

[68] Rapport de la Commission des affaires juridiques du Parlement européen contenant des recommandations à la Commission. Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL)), A. Le rapport traite des cas de la robotique et de l'intelligence artificielle ; il appréhende également - mais de manière moins élaborée -

juridique propre aux robots autonomes les plus sophistiqués.[69] Cela impliquerait que ces machines puissent être tenues responsables des dommages causés, passant ainsi d'une responsabilité du fait d'autrui à une responsabilité directe du robot. Il s'agit en l'occurrence de la thèse du « robot-sujet » de droit.[70]

Cette perspective soulève des questions juridiques délicates, notamment en ce qui concerne l'exercice de professions régulées comme la médecine. La médecine pourrait dès lors être exercée par des machines ou à terme par des robots dotés d'intelligence artificielle. Et comme le suggère Clémentine LEQUILLERIER, « *les actes de diagnostic et de prescription (…) seraient alors entièrement assurés par des machines. Les interventions chirurgicales pourraient, quant à elles, être intégralement réalisées par des robots chirurgicaux dotés d'intelligence artificielle, (ces robots n'auraient pas besoin d'être « pilotés » par des chirurgiens)* »[71]. L'auteur ne manque de renvoyer à l'analyse de Jacques LUCAS qui s'interrogeait - dès 2017 - sur l'hypothèse de voir un jour « *une machine ou un*

---

les cas des véhicules et des drones, ainsi que les robots de soins personnels et les robots médicaux.

[69] Il a été ainsi suggéré « la création, à terme, d'une personnalité juridique spécifique aux robots, pour qu'au moins les robots autonomes les plus sophistiqués puissent être considérés comme des personnes électroniques responsables, tenus de réparer tout dommage causé à un tiers ; il serait envisageable de conférer la personnalité électronique à tout robot qui prend des décisions autonomes ou qui interagit de manière indépendante avec des tiers. ». Résolution du Parlement européen du 16 février 2017 contenant des recommandations à la Commission concernant des règles de droit civil sur la robotique (2015/2103(INL)), 59., f.

[70] « L'inadaptation du droit positif conduit certains à vouloir admettre que le robot devienne un sujet de droit. Par la création d'une nouvelle fiction juridique, le robot serait qualifié de personne et deviendrait titulaire de droits et d'obligations » : MENDOZA-CAMINADE Alexandra, « Le droit confronté à l'intelligence artificielle des robots : vers l'émergence de nouveaux concepts juridiques », D., 2016, p. 445 et suiv.

[71] LEQUILLERIER Clémentine, « L'impact de l'intelligence artificielle sur la santé », op.cit., p.15.

*système de traitement algorithmique de données avec retour de conseils sur la conduite à tenir [puissent] être accusés d'exercice illégal de la médecine (...) ?* »[72] L'on pourrait ainsi admettre qu'un robot doté d'intelligence artificielle puisse réaliser des actes de diagnostic ou de prescription sans intervention humaine. Dans cette hypothèse, les interrogations sur l'exercice illégal de la médecine deviennent pertinentes. Fort heureusement, les voix incarnant cette première option sont peu audibles, car il s'agirait comme l'écrit Grégoire LOISEAU, d'une véritable « monstruosité juridique »[73].

S'inscrivant à l'opposé de la première option, plusieurs juristes plaident en faveur d'une révision des régimes juridiques de responsabilité. Au demeurant, la démarche de révision n'est pas nouvelle, en ce que de nombreuses règles ont déjà fait l'objet d'une adaptation au fil du temps et des évolutions techniques, à l'instar de celles qui régissent l'examen des caractéristiques génétiques d'une personne ou l'utilisation des techniques d'imagerie cérébrale. De manière générale, les défis posés par la complexité et l'autonomie des systèmes d'IA, de même que la multiplicité des acteurs impliqués dans leur développement, rendent la responsabilité très difficile à établir. Cela requiert des ajustements spécifiques[74] afin de garantir une réparation idoine aux victimes.

En effet, l'une des difficultés majeures réside dans le fait qu'un système d'intelligence artificielle puisse être la cause directe

---

[72] LUCAS Jacques, « Les patients et le médecin dans l'écosystème numérique », in. Bulletin de l'Académie Nationale de Médecine, 2017, 201, nos 1-2-3, 121-128, séance du 17 septembre 2017, p.126.

[73] LOISEAU Grégoire, « La personnalité juridique des robots : une monstruosité juridique-Libres propos », in. JCP G2018, 597:<https://gregoireloiseau.fr/2018/05/28/la-personnalite-juridique-des-robots-une-monstruosite-juridique/>

[74] MENDOZA-CAMINADE Alexandra, « Le droit confronté à l'intelligence artificielle des robots : vers l'émergence de nouveaux concepts juridiques », in. D., 2016, p. 445.

ou indirecte d'un dommage, car les activités, les dispositifs ou encore les processus physiques ou virtuels gérés par des systèmes d'IA peuvent techniquement être la cause directe ou indirecte d'un préjudice ou d'un dommage. Cependant, en raison de la connectivité et de l'autonomie des systèmes d'IA, il pourrait, dans la pratique, s'avérer difficile d'identifier les contributions spécifiques des êtres humains ou les décisions de conception qui sont à l'origine d'actes préjudiciables spécifiques commis par des systèmes d'IA.

Nonobstant, et comme semble l'affirmer très tôt la Résolution du Parlement européen en date du 20 octobre 2020, « *il est néanmoins possible de contourner cet obstacle en tenant pour responsables l'ensemble des personnes qui, tout au long de la chaîne de valeur, créent, entretiennent ou contrôlent le risque associé au système d'IA.* »[75] En d'autres termes, il est souvent nécessaire de remonter à la personne ayant conçu ou déployé ce système pour établir cette responsabilité. Il est essentiel d'identifier toutes les personnes qui, à chaque étape de la chaîne de valeur, créent, entretiennent ou contrôlent ce risque.

C'est donc la seconde option qui a été privilégiée, en l'occurrence celle de la modernisation, si ce n'est le renforcement des régimes de responsabilité objective. Concrètement, il s'agira de revoir les règles juridiques applicables en matière de responsabilité du fait des produits défectueux. Au sens juridique, les produits défectueux constituent une catégorie juridique[76] de

[75] Point 6. Résolution du Parlement européen du 20 octobre 2020 contenant des recommandations à la Commission sur un régime de responsabilité civile pour l'intelligence artificielle (2020/2014(INL))
JO C 404 du 6.10.2021, pp. 107-128.

[76] « (…) Les catégories juridiques constituent (…) un facteur de simplification du droit. » : BERGEL Jean-Louis, Théorie générale du droit, 5e éd., Dalloz, 2012, p. 238, §180.

choses différentes des autres, et partant, le régime de responsabilité y afférent l'est également. Le législateur tunisien sera sans doute appelé - sinon tenu - de le faire très rapidement. Il conviendra de procéder à une révision des règles actuellement en vigueur en matière de responsabilité objective des fabricants pour les produits défectueux, allant des technologies intelligentes aux produits pharmaceutiques. L'impératif de révision est double. Il s'agit d'améliorer la sécurité juridique des entreprises innovantes qui travaillent dans le domaine des systèmes IA, et de garantir, dans le même temps, une indemnisation équitable des victimes d'un dommage, si elles démontrent un défaut du système IA. Les systèmes d'IA étant assimilés à des logiciels, la révision des règles de responsabilité du fait des produits défectueux permettra d'élargir le champ d'application à tous les produits qui sont des systèmes d'IA ou dont au moins un des composants est un système d'IA. L'Union européenne l'a fait à travers l'adoption de Directive 2024/2853 du 23 octobre 2024 relative à la responsabilité du fait des produits défectueux qui vient d'être publiée le 18 novembre au Journal officiel de l'UE. Cette directive s'applique ainsi aux systèmes d'Intelligence artificielle[77] ; elle a

[77] Le nouveau texte abroge et remplace la directive n° 85/374/CEE du Conseil du 25 juillet 1985 qui avait été transposée dans le code civil aux articles 1245 à 1245-17 (ancien articles 1386-1 à 1386-18). Or, malheureusement ce texte était limité aux seuls biens meubles (art. 1245-2 code civil) ce qui excluait de facto les logiciels et par voie de conséquence, les systèmes d'IA ainsi que les évolutions technologiques récentes. Cette nouvelle donne réglementaire doit être donc analysée en complément du Règlement européen sur l'intelligence artificielle (AI Act), publié au Journal Officiel de l'UE le 12 juillet 2024. Ce texte d'application directe dans les systèmes juridiques de l'UE suit une approche par les risques. Il vise à garantir, au sein des États membres, l'utilisation des systèmes d'IA de façon éthique, sécurisée et fidèle aux droits fondamentaux (voir également la convention-cadre du Conseil de l'Europe sur l'intelligence artificielle du 5 septembre 2024). Les États membres doivent transposer la directive dans leur droit national dans les deux ans, soit 20 jours après sa publication au JOUE (c'est-à-dire le 9 décembre 2026). Il va sans le dire que les entreprises impactées par le nouveau texte

notamment pour objectif de réparer les dommages causés par lesdits systèmes. C'est à dessein que la notion de produit de l'article 4 de la Directive se trouve significativement élargi afin d'inclure les logiciels, y compris les systèmes d'IA tels qu'ils ont été définis dans le cadre du Règlement européen relatif à l'intelligence artificielle (AI Act).[78] L'article 8 concerne la responsabilité des opérateurs économiques en cas de produits défectueux (fabricant du produit ou d'un composant, importateur, mandataire ou prestataire de services d'exécution, voire toute personne qui modifie le produit de manière substantielle en dehors du contrôle du fabricant). L'article 9 pose, quant à lui, des protocoles spécifiques concernant la divulgation des éléments de preuve. Enfin, l'article 10 établit une présomption de défectuosité du produit, ce qui implique d'engager la responsabilité des fabricants de produits ou de services se fondant sur l'IA en cas de défectuosité du système d'IA intégré au produit ou encore au service de santé.

Concernant la question du défaut et du lien de causalité, il est effectivement plus aisé de définir le caractère défectueux d'un produit, sans nécessairement lui imputer un comportement fautif, et ce, *« indépendamment de toute faute.* »[79] La qualification de la faute requise pour engager la responsabilité constitue une source de complexité notable, en ce que les systèmes d'intelligence

(spécialement les fabricants et les acheteurs) seront bien avisées d'anticiper la transposition dans le code civil en revisitant leurs clauses contractuelles ainsi que le périmètre de couverture de leurs polices d'assurance.

[78] La directive définit le terme « produit » comme « tout meuble, même s'il est incorporé dans un autre meuble ou dans un immeuble ou interconnecté avec celui-ci ; le terme comprend l'électricité, les fichiers de fabrication numériques, les matières premières et les logiciels. ».

[79] Voir notamment l'article de CHRISTODOULOU Hélène, « La responsabilité civile extracontractuelle à l'épreuve de l'intelligence artificielle », in. Lexbase Droit privé, 2019, n°807, pp. 04-05.

artificielle, de par leur nature autonome et leur fonctionnement algorithmique, sont susceptibles de causer un préjudice sans qu'une faute caractérisée, au sens classique du terme, puisse être nécessairement imputée à un acteur identifiable.

De nouveaux critères d'imputabilité sont introduits pour s'adapter à des produits de l'envergure des systèmes « IA ». Il sera ainsi nécessaire de vérifier la nature du défaut : « (…) *Pour les modèles employés par les « IA », le défaut serait à apprécier s'il se produit de manière systémique par comparaison avec les normes techniques existantes dans le secteur concerné ou, en l'absence de normes, probablement avec les performances humaines.* »[80] Il s'agira d'une présomption simple de défectuosité si le défendeur n'est pas en mesure de fournir des preuves[81]. Pour Yannick MENECEUR, cette présomption « *peut être établie si le demandeur démontre que le produit ne remplit pas les exigences de sécurité obligatoires (imposées par le droit de l'Union ou le droit national), ou, alternativement, s'il démontre que les dommages sont causés par un dysfonctionnement manifeste du produit résultant de son utilisation dans des circonstances normales.* »[82] La responsabilité s'appliquera donc lorsque des preuves de défauts dans les systèmes d'intelligence artificielle seront établies, mais cela exige une réelle expertise technique.

---

[80] Résolution du Parlement européen du 20 octobre 2020 comportant des recommandations à la Commission sur un régime de responsabilité civile pour l'intelligence artificielle (2020/2014(INL)), *JO C 404 du 6.10.2021, p. 107–128,* Point 8 : « La directive sur la responsabilité du fait des produits a, depuis plus de 30 ans, prouvé son efficacité en tant qu'outil permettant d'obtenir réparation du préjudice causé par un produit défectueux, mais il convient néanmoins de la réviser (…) ».

[81] C'est le cas de la Directive sur la responsabilité du fait des produits.

[82] MENECEUR Yannick, « Les nuances des nouveaux régimes de responsabilité des IA en Europe », in. Actualités Juridiques, 15/03/2023, consultable sur : <https://www.actu-juridique.fr/id/AJU356431>

D'aucuns proposent carrément l'application des règles de la responsabilité médicale. Le principe en la matière - rappelons-le - est celui d'une responsabilité pour faute, à moins que le dommage n'ait été causé par un défaut d'un produit de santé. Il convient de rappeler à cet égard que les produits de santé ne sont pas seulement les médicaments mais également les dispositifs médicaux, et de manière plus large, tout instrument, appareil, équipement, matière, produit destiné par le fabricant à être utilisé chez l'homme à des fins médicales.[83] Il en est également de même pour tout logiciel destiné par le fabricant à être utilisé spécifiquement à des fins diagnostiques ou thérapeutiques. Dans ces conditions, l'on peut considérer que le système d'intelligence artificielle conçu à des fins diagnostiques ou thérapeutiques constitue un produit de santé. Il y a lieu de craindre, tout de même, un abus du recours à la responsabilité médicale, c'est-à-dire, la crainte que celle-ci ne soit invoquée à outrance, et en toute instance. Pour autant, la responsabilité des professionnels de santé « *ne peut être engagée que si le système présente un défaut. Or, non seulement tout dommage n'est pas nécessairement le fait d'un défaut, mais la preuve de celui-ci peut être en pratique extrêmement difficile à rapporter*. »[84]

Pour revenir à la question de la responsabilité objective, les dommages pouvant faire l'objet d'une indemnisation sur la base de ladite responsabilité, notamment dans le cas des produits défectueux, sont très larges et peuvent se rapporter « *tant aux*

[83] Pour une définition exhaustive, voir SERRE Marie-Paule et WALLET-WODKA Deborah, « Chapitre 1. Qu'est-ce qu'un produit de santé ? Définition et typologie », in. Marketing des produits de santé, Stratégies d'accès au marché, Médicaments remboursables, selfcare, cosmétiques et aliments santé, Dunod, 2e éd., 2014, pp.15-37.
[84] LOISEAU Grégoire, « Intelligence artificielle et santé », op. cit., p.22.

*dommages corporels et matériels qu'aux dommages résultant de la perte de données.* »[85]

Nous avons évoqué tantôt la responsabilité en cas d'un système d'IA défectueux, et ce, lorsque ledit système est considéré comme un produit de santé. De par une jurisprudence constante, l'implication d'un produit de santé dont un défaut est établi engage la responsabilité sans faute du professionnel de santé. La responsabilité du fait des produits défectueux pourra certainement être invoquée lorsque le dommage est causé par un système d'intelligence artificielle utilisé à des fins médicales.

Appréhendons à présent la responsabilité civile pour faute. Dans la plupart des cas, les règles applicables assurent une protection adéquate aux personnes lésées par des tiers perturbateurs, et notamment des pirates informatiques, sur la base d'une action fondée sur une faute. D'aucuns considèrent que les personnes lésées devraient également bénéficier d'une présomption de faute de l'opérateur qui pourra éventuellement s'exonérer de sa responsabilité en apportant la preuve du respect du devoir de diligence. Il va sans dire que l'opérateur/exploitant d'un système d'IA défectueux peut, après avoir intégralement indemnisé la victime du préjudice ou des dommages causés, se retourner contre le producteur du système IA défectueux, c'est à dire le poursuivre en justice conformément aux règles juridiques en vigueur applicables en matière de responsabilité du fait des produits défectueux.

Cependant, l'application des règles de la responsabilité soulève une difficulté. S'il est vrai que les règles classiques

[85] MENECEUR Yannick, « Les nuances des nouveaux régimes de responsabilité des IA en Europe », op. cit.

demeurent applicables[86], en pratique, la condition relative à l'identification du lien de causalité entre le dommage et la faute s'avère difficile à vérifier. En effet, s'agissant des systèmes d'IA, relier les dommages à l'erreur humaine est tout à fait complexe, en raison, d'une part, de la diversité des acteurs (fournisseurs, importateurs, utilisateurs, etc.), et d'autre part, de la difficulté inhérente à la traçabilité. L'exemple la chirurgie robotisée, qui permet aux chirurgiens de réaliser avec succès certaines interventions chirurgicales complexes où la dextérité de la main humaine est insuffisante, illustre cette difficulté. Les robots chirurgicaux sont contrôlés par des algorithmes lesquels peuvent être entachés de biais.[87] L'on peut se référer ici à la proposition de Directive[88] sur la responsabilité civile extra-contractuelle en matière d'IA en date du 28 septembre 2022. La personne lésée à la suite de dommages liés au SIA est en droit d'obtenir réparation à condition de pouvoir apporter la preuve de la faute commise par le fournisseur du SIA. Cependant, l'interposition d'une personne

[86] Voir VINEY Geneviève, JOURDAIN Patrice (avec la collaboration de CARVAL Suzanne), GHESTIN Jacques (dir.), Traité de droit civil, Les conditions de la responsabilité, LGDJ, 4e éd., 2013, n° 256 et suiv.

[87] Voir notamment GRUSON David, « Régulation positive de l'intelligence artificielle en santé : les avancées de la garantie humaine algorithmique », in. Dalloz IP/IT, n°3, 2020, pp.165-170.

[88] Version Consolidée du Traité sur le Fonctionnement de l'Union Européenne - Sixième Partie - Dispositions Institutionnelles et Financières - Titre I - Dispositions Institutionnelles - Chapitre 2 - Actes Juridiques de l'Union, Procédures d'adoption et autres Dispositions - Section 1 - Les Actes Juridiques de l'Union Article 288 (Ex-Article 249 TCE) : « Pour exercer les compétences de l'Union, les institutions adoptent des règlements, des directives, des décisions, des recommandations et des avis.
Le règlement a une portée générale. Il est obligatoire dans tous ses éléments et il est directement applicable dans tout État membre.
La directive lie tout État membre destinataire quant au résultat à atteindre, tout en laissant aux instances nationales la compétence quant à la forme et aux moyens.
La décision est obligatoire dans tous ses éléments. Lorsqu'elle désigne des destinataires, elle n'est obligatoire que pour ceux-ci.
Les recommandations et les avis ne lient pas. »

entre le fonctionnement du SIA et le préjudice subi fait écran et impliquera l'application du régime de responsabilité de droit commun.

De manière plus concise, nous dirons que la mise en œuvre de la responsabilité peut être scindée ainsi : avant, pendant et après le déploiement d'un système d'IA. Avant le déploiement, le système basé sur l'IA doit être conçu de telle sorte que la validité de son évaluation puisse être vérifiée, qu'il soit fiable lorsqu'il est nourri avec les mêmes données, et qu'il soit lui-même soumis à une analyse d'impact. Pendant son déploiement, le système fondé sur l'IA doit faire l'objet d'un suivi et d'un audit continus réalisés par une structure mise en place par les utilisateurs du système d'IA ou ses « créateurs ». Cela devrait aussi inclure la possibilité pour les personnes soumises au système d'IA de faire appel du résultat de l'évaluation. Après le déploiement d'un système basé sur l'IA, des mécanismes de régulation devraient être mis en place afin – le cas échéant - de rectifier les décisions considérées potentiellement « injustes » prises par le système basé sur l'IA, et ce, outre la responsabilité pesant sur les opérateurs en cas de survenance d'un dommage lié à l'usage du système d'IA.

Ceci étant précisé, il convient de rappeler, qu'en principe, les médecins – et de manière plus générale les prestataires de soins de santé - sont responsables en dernier ressort des traitements qu'ils administrent. Dans les cas de faute professionnelle, ce sont les prestataires qui doivent justifier leurs décisions. Vous l'aurez certainement compris, il ne s'agit pas de transférer la responsabilité du traitement administré aux développeurs du système d'IA ; ces derniers ne sauraient aucunement se substituer au médecin pour assumer la

responsabilité liée à une faute professionnelle. Mais il est probablement trop tôt pour se prononcer, car les systèmes d'IA ne possèdent pas actuellement les capacités technologiques suffisamment avancées pour reproduire les tâches que les médecins effectuent. Les traitements sont souvent très individualisés, ce qui ne cadre guère avec les « compétences » de l'IA plutôt versée dans les tâches à forte répétition et à faible risque. L'intégration de l'IA dans les processus de traitement médical semble peu probable dans un avenir proche.

Une chose demeure certaine, tous les opérateurs des systèmes d'IA à haut risque devraient souscrire une assurance responsabilité civile. Le régime d'assurance obligatoire pour le système d'IA à haut risque devrait couvrir l'étendue de l'indemnisation sans pour autant que les primes d'assurance ne soient élevées, car il risque de se transformer en un obstacle, voire « *un frein à la recherche et à l'innovation.* »[89]Enfin, nous n'évoquerons pas - dans le cadre de cette intervention - la responsabilité pénale même si l'on sait que, aujourd'hui et plus encore demain, que les systèmes d'IA peuvent être utilisés, non seulement pour diagnostiquer des pathologies, mais pourraient à terme être utilisés pour améliorer les capacités et performances des individus, et partant pour « *constituer des classes humaines supérieures, génétiquement augmentées*. »[90]

[89] Introduction, Point 24 de la Résolution du Parlement européen du 20 octobre 2020 contenant des recommandations à la Commission sur un régime de responsabilité civile pour l'intelligence artificielle (2020/2014(INL)), P9_TA(2020)0276, Un régime de responsabilité civile pour l'intelligence artificielle : <https://www.europarl.europa.eu/doceo/document/TA-9-2020-0276_FR.html>

[90] Grégoire LOISEAU, « Intelligence artificielle et santé », op. cit., p.23.

## Conclusion

L’intelligence artificielle appliquée à la médecine fait appel à une multitude de règles juridiques, et le risque d’enchevêtrement est bien présent. La doctrine s’accorde – pour l’heure – à considérer qu’il n’y a pas lieu de lui attribuer une personnalité juridique autonome en termes de droits et obligations. La thèse qui prévaut milite plutôt en faveur de l’application du régime juridique de la responsabilité du fait des produits défectueux et du régime de la responsabilité objective.

L'intégration croissante de l'IA dans les processus décisionnels médicaux soulève également des interrogations quant à la responsabilité professionnelle. À quel point peut-on déléguer des décisions médicales à une machine sans compromettre le principe fondamental de la responsabilité humaine ? Les médecins doivent rester au cœur du processus décisionnel, même lorsque les systèmes d'IA sont utilisés pour appuyer ou corroborer leurs choix cliniques. Cela implique une réflexion approfondie sur le rôle que doit jouer l'humain face à l'automatisation croissante des soins.

# Rapport de synthèse

*Wafa HARRAR MASMOUDI, Rym RAFRAFI*

Le choix de la thématique de la conférence annuelle du CNEM qui coïncide avec le 30ème anniversaire de la création du Comité national d'éthique médicale intitulée « L'intelligence artificielle en santé - Enjeux éthiques » est parfaitement fondée et justifiée. Ce choix se fonde en effet sur la nécessité impérieuse d'explorer les implications multidimensionnelles de l'intégration de l'IA dans le secteur de la santé. Cette thématique est d'une actualité brûlante, tant l'IA transforme les pratiques médicales. L'ambition de cette conférence annuelle a donc été d'explorer les diverses dimensions de l'intégration de l'intelligence artificielle (IA) dans le secteur de la santé mais également d'offrir un espace de réflexion et d'échange sur ces enjeux multidimensionnels, visant à éclairer les perspectives futures de l'intelligence artificielle dans le domaine de la santé.

C'est ainsi que les communications ont successivement portées sur les aspects médicaux, éthiques, juridiques, et pédagogiques que soulève l'utilisation de l'IA dans le domaine de la santé.

La première communication « L'IA et santé : apports et risques » a été présentée par madame Saoussen HANTOUS, professeur de médecine à la Faculté de Médecine de Tunis, et chef de service de Radiologie à l'hôpital Abderrahman Mami. Professeur Saoussen HANTOUS a posé les premiers jalons en expliquant comment l'intelligence artificielle (IA) est-elle en train de transformer le paysage de la santé, offrant des opportunités

sans précédent pour la médecine tout en soulevant des préoccupations éthiques et pratiques.

Professeur HANTOUS perçoit l'intelligence artificielle (IA) comme un acteur clef dans la transformation du secteur de la santé. Les apports de l'IA sont en effet indéniables allant de l'aide au diagnostic puisque l'IA facilite la détection précoce de diverses pathologies, notamment en imagerie médicale, ainsi que dans des spécialités telles que la dermatologie et l'ophtalmologie, à la personnalisation des traitements[91], sans occulter le volet préventif, en ce que l'IA est également reconnue pour sa capacité à anticiper les épidémies et à surveiller les effets secondaires des médicaments, contribuant ainsi à une meilleure prévention sanitaire.

Professeur HANTOUS a de surcroît mis en exergue l'apport de l'IA en matière de suivi à distance et à temps réel des patients, ce qui est à même de compenser le manque croissant de personnel médical, de même que les avancées réalisées grâce à l'IA en chirurgie, et les perspectives hautement encourageantes attenantes au volet de la formation médicale.

Cependant - insiste le professeur HANTOUS - ces apports (pour ne pas parler de bénéfices) ne doivent pas occulter les risques associés à l'intégration de l'IA en matière de soins. En effet, et bien qu'elle promette des avancées significatives, l'IA présente également des risques notables. En premier lieu, les biais algorithmiques peuvent engendrer des erreurs diagnostiques, compromettant ainsi la sécurité des patients. De surcroît, la sécurité informatique est menacée par des attaques qui ne manquent de mettre en péril la fiabilité des systèmes d'IA.

[91] L'IA permet un suivi individualisé des maladies chroniques et favorise l'élaboration de traitements ciblés basés sur l'analyse des données génomiques.

La protection des données personnelles soulève également des préoccupations, s'agissant notamment de l'utilisation non autorisée de données sensibles. En outre, la dépendance croissante aux algorithmes risque d'altérer le libre arbitre du médecin et du patient, tout en obscurcissant la question de la responsabilité légale en cas d'erreur algorithmique. Par ailleurs, l'essor technologique pourrait nuire à l'empathie essentielle à la relation soignant-soigné et affaiblir les compétences critiques des professionnels de santé. Enfin, rappelle le professeur HANTOUS, il convient de noter que le financement inégal des technologies d'IA risque d'accentuer les disparités dans l'accès aux soins. Ainsi, une régulation appropriée et une réflexion éthique approfondie sont indispensables pour garantir que l'IA serve effectivement le bien-être des patients tout en préservant les valeurs fondamentales de la médecine.

La seconde communication est originale à plus d'un titre : l'interaction novatrice entre un professeur de philosophie, monsieur Youssef BEN OTHMAN et une professeur de médecine, madame Rym RAFRAFI, pour appréhender l'Intelligence Artificielle à l'épreuve de l'éthique a scellé une alliance complice invitant à une réflexion profonde. En effet, la collaboration entre un professeur de philosophie et une professeure de médecine a donné lieu à une synergie intellectuelle remarquable favorisant un dialogue interdisciplinaire qui invite à une réflexion approfondie sur les implications morales et éthiques des technologies émergentes. En conjuguant les perspectives philosophiques sur la nature humaine, la responsabilité et la justice avec les réalités pratiques et les défis posés par l'IA dans le domaine médical, l'intervention du professeur RAFRAFI et du professeur BEN OTHMAN a ouvert la voie à une compréhension nuancée

des enjeux posés par l'IA. Celle-ci impose désormais une réflexion profonde sur ce que signifie l'être humain dans un monde où les machines peuvent simuler des comportements intelligents.

L'intervention s'est articulée, en premier, autour des sources philosophiques : « *une vision mécaniste du monde et de l'homme* » en vue de mettre en exergue les enjeux éthiques et philosophiques posés par l'IA. La modernité scientifique, avec son penchant pour le mécanicisme, a marqué la naissance de l'Homme moderne, ce qui représente une transition significative dans la pensée philosophique. Les conférenciers ont rappelé que René Descartes a introduit la dualité âme/corps, posant que l'âme est distincte du corps, tout en développant la thèse de l'animal-machine laquelle considère les animaux comme des automates dépourvus de conscience. En renvoyant aux travaux de Julien Offray de La Mettrie (un saut dans le temps, car un siècle plus tard), les conférenciers se sont arrêtés *a fortiori* sur l'idée de l'Homme-machine intégrant une vision matérialiste qui remet en question les notions traditionnelles de l'âme. C'est cette conception qui a ouvert la voie à des réflexions contemporaines sur l'intelligence artificielle, suggérant que les avancées technologiques représentent une mutation technoscientifique fondamentale.

Les conférenciers ont en outre établi un constat, à savoir, celui des défis cruciaux que pose l'émergence de l'intelligence artificielle aux sciences humaines et sociales, et du désormais inéluctable débat autour de la notion d' « *Homo numericus* », qui désigne l'individu contemporain façonné par le numérique et ses implications sur l'identité humaine

Le troisième axe de l'intervention s'est articulé autour du postulat de la rupture épistémologique. L'IA semble engendrer une véritable « rupture épistémologique », un changement radical dans la manière dont les connaissances sont construites et comprises. En introduisant des outils d'analyse automatisés, l'IA remet en question les paradigmes traditionnels.

Les conférenciers ont passé en revue les nombreux et complexes défis éthiques. L'émergence des algorithmes autonomes - affirment-ils - soulève des questions cruciales relatives à l'équité et aux biais inhérents qui peuvent engendrer des conclusions erronées et discriminatoires, illustrant ainsi la nécessité impérative de transparence dans leur fonctionnement. Dans des domaines sensibles tels que la santé, la justice et les transports, la question de la responsabilité a émergé de manière prépondérante : qui porte la charge des erreurs commises par l'intelligence artificielle ? En parallèle, l'explicabilité des décisions complexes prises par ces systèmes devient essentielle pour garantir une compréhension claire et accessible, alors même que l'agentivité de l'IA, dépourvue de conscience et d'intention humaine, pose des défis éthiques majeurs. Les enjeux de la vie privée et de la surveillance, exacerbés par l'utilisation extensive des données, soulèvent des préoccupations quant aux menaces potentielles qui pèsent sur l'intimité individuelle. Dans le contexte économique, l'automatisation présente des risques significatifs de perte d'emplois, appelant à une transition vers de nouvelles compétences et à une clarification des choix technologiques. Enfin, les enjeux philosophiques de l'identité et de la déshumanisation se posent face à une IA opaque, interrogeant les limites de cette technologie par rapport à l'humain et engageant

une réflexion sur les implications éthiques du transhumanisme et des augmentations technologiques.

La transition est parfaitement assurée puisque la troisième intervention a justement appréhendé « *L'Humanisme, La Déshumanisation et l'IA* ». Madame Salma DEROUICHE EL KAMEL enseignante-chercheuse en psychologie clinique à l'Université de Tunis, a souligné que l'intelligence artificielle (IA) a transformé profondément les pratiques en psychologie et en santé mentale, nécessitant ainsi une réévaluation des méthodes cliniques au-delà de ces domaines, notamment en médecine générale et spécialisée.

La conférencière a posé la question cruciale de savoir si, dans cette évolution technologique, l'humanité ne risque-t-elle pas de perdre une part essentielle de son essence. Il est ainsi essentiel d'établir des limites à ne pas franchir et d'identifier les responsabilités liées à la mise en place de garde-fous contre l'excès d'innovation. Elle a précisé en outre qu'il convient de parler d'"IAs" au pluriel, afin de rendre compte de la diversité des applications de ces technologies, allant du diagnostic à la thérapie en passant par l'analyse des données. Madame Salma DEROUICHE EL KAMEL a évoqué l'effervescence actuelle entourant les IAs, attribuée à plusieurs facteurs, tels que les avancées rapides en informatique et en apprentissage automatique, l'impact des crises humanitaires récentes - en particulier celle du COVID-19 - et la large diffusion des technologies, illustrée par des agents conversationnels comme ChatGPT.

S'agissant de la santé mentale, la conférencière a mis en avant l'existence des robots conversationnels en tant qu'outils particulièrement humanisés. Elle a cité des recherches prometteuses, comme celle relative au coach vocal virtuel

« Lumen », qui a montré des résultats positifs dans le traitement de l'anxiété et de la dépression. De même, le chatbot "DEPRA", dédié au dépistage précoce de la dépression, et qui vise à atteindre des populations souvent négligées pour des raisons variées.

Cependant, Madame Salma DEROUICHE EL KAMEL a exprimé des craintes légitimes entourant la protection des données personnelles, la fiabilité des diagnostics, et les biais algorithmiques. Elle a mentionné que l'*American Psychological Association* a formulé des principes éthiques pour encadrer l'intégration de l'IA dans le domaine de la santé mentale, en insistant sur l'importance de centrer la technologie sur l'humain, de valoriser la diversité, et de mener des recherches approfondies. Elle a ensuite abordé la question de la déshumanisation des pratiques cliniques, se demandant si l'IA pourrait un jour remplacer les interactions humaines essentielles. Selon ses propos, bien que les robots et les logiciels puissent soutenir les praticiens, ils ne sauraient se substituer à la dimension humaine de l'empathie et des échanges intersubjectifs. Elle a averti que l'éventualité d'une IA consciente à l'avenir pourrait nécessiter une préparation sociétale face à ses implications éthiques.

Enfin, la conférencière a conclu que le débat actuel s'inscrit dans une réflexion plus large sur l'évolution humaine et les comportements, où le désir d'immortalité est constamment alimenté par les avancées technologiques. Elle a souligné que cette quête de toute-puissance, soutenue par les progrès de l'IA, interroge en réalité notre rapport à la mort et à la vie, mettant en lumière une tension entre le désir de vivre et l'inévitabilité de la mort.

Après une pause gourmande bien méritée et grandement appréciée, la seconde séance a été consacrée à la régulation de l'IA dans le domaine de la santé. Par la suite, la parole a été donnée aux jeunes professionnels de la santé, dont les perspectives étaient essentielles pour enrichir le débat. Ils ont partagé leurs réflexions sur les bénéfices potentiels de l'IA dans leur pratique, ainsi que sur les dérives et les défis qui pourraient en découler, apportant ainsi un éclairage contemporain sur l'intégration de ces technologies dans le paysage de la santé moderne.

La quatrième intervention de la Journée a été consacrée au volet juridique. Madame Wafa HARRAR MASMOUDI, Professeur de Droit à la Faculté des sciences juridiques, politiques et sociales - Université de Carthage, a commencé par rappeler que l'impact de l'intelligence artificielle (IA) dans le domaine de la santé n'est plus à démontrer et que le cadre juridique y afférent est en pleine effervescence. Elle a passé en revue les initiatives visant à encadrer l'utilisation de l'IA aux États-Unis et en Europe, et s'est attardée sur la définition apportée par le seul instrument juridique à caractère contraignant, en l'occurrence le Règlement européen, (Loi sur l'IA) qui est entré en vigueur le 1er août 2024. Ladite législation - a précisé la conférencière - classe les systèmes IA selon leur niveau de dangerosité, allant du risque inacceptable, au risque élevé, et enfin au risque minimal.

Elle a ainsi examiné en premier les obligations découlant de l'usage des systèmes d'IA en matière de santé, qui sont modulées en fonction desdits risques avant d'identifier les règles applicables en matière de responsabilité, une responsabilité qui demeure difficile à circonscrire. La conférencière a expliqué comment les obligations des opérateurs d'IA en médecine

dépendent des risques identifiés. Ces obligations incluent des exigences de transparence, d'explicabilité et d'information des patients. Professeur HARRAR MASMOUDI a également souligné l'importance de la protection des données personnelles et du respect du consentement éclairé, qui doit s'adapter aux nouvelles réalités technologiques.

La question de la responsabilité en cas de préjudice causé par l'IA a été longuement analysée. Deux options majeures se détachent. Il s'agit soit d'attribuer une personnalité juridique à l'IA, soit d'appliquer les régimes de responsabilité classiques. La tendance s'oriente - affirme la conférencière - vers une adaptation des règles existantes, afin de garantir que les victimes puissent obtenir réparation. Elle a insisté sur la nécessité de réviser les régimes de responsabilité en raison des risques inéluctables inhérents à l'utilisation des systèmes intégrés dans les dispositifs médicaux, tout en assurant la sécurité juridique en matière d'innovation. Le législateur tunisien, à l'instar de plusieurs législateurs à travers le monde, sera nul doute appelé à engager une réforme substantielle afin de traduire juridiquement les réalités imposées par l'IA. Cette réforme doit s'inscrire dans une optique d'amélioration du volet attenant à la sécurité juridique pour les entreprises innovantes tout en garantissant une juste indemnisation pour les victimes.

Enfin, la voix des jeunes professionnels de la santé a été revigorante. L'exploration de l'intégration de l'IA dans les « Pratiques des Résidents en Médecine » à travers une enquête nationale diligentée par Docteur Houssem ABIDA, de l'Association des jeunes internistes tunisiens, a pour objectif l'analyse de la perception par les résidents, de l'implémentation de l'IA dans la pratique hospitalière et l'enseignement académique.

L'importance de la représentativité des spécialités et la mise en lumière des deux principaux aspects du parcours des résidents constituent les points forts de l'enquête. Les résultats indiquent que la nouvelle génération de médecins est largement influencée par l'IA, avec un intérêt croissant pour des méthodes d'apprentissage facilitées par cette technologie. Les jeunes résidents se montrent ouverts à l'intégration de l'IA dans le diagnostic et la prise de décisions thérapeutiques.

Cependant, plusieurs limites ont été identifiées. Il appert ainsi que la taille relativement faible de l'échantillon peut affecter la généralisation des résultats, et que l'absence d'évaluation des connaissances des résidents sur l'IA représente une lacune de taille pour assurer une compréhension approfondie de leurs perceptions. Le questionnaire révèle que la génération actuelle de résidents est non seulement imprégnée par l'IA, mais également enthousiaste à l'idée de l'utiliser dans leur pratique médicale. L'adoption d'un cadre législatif adapté est désormais essentielle pour concilier entre l'innovation et la protection des patients, tout en répondant aux défis juridiques posés par ces nouvelles technologies.

En guise de conclusion, il est permis d'affirmer que cette exploration multidimensionnelle des enjeux liés à l'intelligence artificielle en matière de santé a réussi à mettre en lumière non seulement ses potentialités mais également ses défis éthiques et sociaux. Il est impératif que cette réflexion soit poursuivie afin de garantir que le développement technologique puisse s'inscrire dans un cadre respectueux des valeurs humaines fondamentales.