

RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

Dire oui et non aux outils d'intelligence artificielle pour préserver la relation de soin

Laurent, Nathanael; Giorgi, Federico; Guillermin, Mathieu

Published in:
Ethica Clinica

Publication date:
2024

Document Version
Version créée dans le cadre du processus de publication ; mise en page de l'éditeur ; généralement non rendue publique

[Link to publication](#)

Citation for pulished version (HARVARD):

Laurent, N, Giorgi, F & Guillermin, M 2024, 'Dire oui et non aux outils d'intelligence artificielle pour préserver la relation de soin', *Ethica Clinica*, vol. 113, pp. 11-25.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Paru dans la [revue francophone d'éthique *Ethica Clinica*](#) : trimestriel 113/2024.

Version finale corrigée par les auteurs.

Dire oui et non aux outils d'intelligence artificielle pour préserver la relation de soin.

Nathanaël LAURENT, Federico GIORGI et Mathieu GUILLERMIN

Résumé :

Depuis quelques années, nous observons une nette accélération de l'introduction d'outils d'intelligence artificielle (IA) de plus en plus puissants dans la plupart de nos environnements tant privés que professionnels. A une époque où la cohabitation des êtres vivants sur notre planète est en crise, l'arrivée aussi massive de techniques imitant et prolongeant certaines capacités cognitives humaines a de quoi poser question. Dans un domaine comme celui de la médecine, l'utilité de certaines IA ne fait certes aucun doute ; mais c'est leur impressionnante performance qui attise justement nos interrogations : ces techniques resteront-elles au service de l'humain ou ce dernier se laissera-t-il impressionner au point de se soumettre aux exigences des IA ? Plus fondamentalement, la relation médecin-patient pourra-t-elle être préservée si le premier laisse son esprit être envahi par la logique calculatrice, et le second se voit remplacé par son image numérisée ?

Pour nous aider à nous orienter face à ces transformations rapides comportant de multiples enjeux, nous proposons une réflexion partant de notre rapport à la technique. Nous mettrons en évidence une posture consistant non pas à opposer le « pour » et le « contre », mais plutôt à dire à la fois oui et non aux outils d'IA. Nous donnons ensuite un aperçu de ce que sont ces outils et de quelques-unes de leurs applications dans le domaine de la santé. Enfin, nous nous tournons vers une illustration tirée d'un projet de recherche-action qui s'intitule « Nouvel Humanisme au temps des Neurosciences et de l'Intelligence Artificielle » (NHNAI). La conclusion récapitule les apports d'un tel parcours réflexif et revient sur sa dimension éthique.

Mots-clés : Intelligence artificielle, médecine, technique, éthique, humain.

Introduction.

Toute nouveauté technologique entraîne de l'engouement et des espoirs, mais aussi en même temps de l'hésitation, voire de la méfiance. Nous ne savons tout simplement pas quelles seront toutes les conséquences de ses applications, et nul ne peut s'assurer entièrement à l'avance de sa sécurité.

En ce qui concerne les techniques dites d' « intelligence artificielle » (IA), reconnaissons qu'elles sont déjà entrées dans notre vie quotidienne sans que nous ne le remarquions. En usant des dernières technologies de l'information et de la communication (smartphones, ordinateurs et autres objets connectés), nous disons donc implicitement « oui » aux IA.

Dans le domaine de la santé, comme dans tout autre, les applications de l'IA en phase de développement sont de plus en plus nombreuses, et certaines sont déjà introduites au cœur même des lieux et pratiques de soins. Ici aussi il peut sembler que nous disions « oui » aux IA. Comment pourrait-il en être autrement ? L'humanité a-t-elle au cours de son histoire pu délibérer au sujet de l'acceptation ou non d'une nouvelle technique ?

D'un point de vue global, le progrès technique fait bel et bien partie de notre histoire. Il nous paraît difficile de donner un exemple de nouvelle technologie qui, sous certaines formes au moins, n'ait pas été adoptée et propagée au sein des sociétés humaines depuis les premiers outils, la maîtrise du feu, mais aussi les techniques d'élevage et de culture, sans oublier les méthodes de soins. Mais nous pourrions *a contrario* toujours donner des exemples de communautés ou d'individus qui résistent et refusent certaines nouvelles technologies. Qu'il s'agisse des Amish (minorité religieuse implantée aux États-Unis et issue du mouvement anabaptiste), des néo-luddiques (en référence à Ned Ludd, un ouvrier anglais qui a protesté contre l'usage des machines à tisser à la fin du 18^{ème} siècle), ou d'autres groupes et individus revendicateurs ou anonymes, dire « non » à certaines technologies fait aussi partie du paysage bigarré de notre communauté humaine.

Face à l'inéluctable arrivée des applications de l'IA dans les pratiques médicales, nous souhaitons dans cet article partir d'une réflexion plus générale portant sur notre rapport à la technique. Ce que nous essaierons ainsi de mettre en évidence, c'est une posture consistant non pas à opposer les tenants du « oui » à ceux du « non », mais consistant à dire à la fois « oui » et « non » aux apports technologiques. Nous pensons que pour préserver le caractère humain indispensable et irremplaçable de toute relation de soin, il est important d'éclairer notre devoir d'accepter et de refuser ce qui est produit par l'être humain à des fins utilitaires.

Dans un premier temps nous présenterons notre position, qui peut à première vue paraître paradoxale, en partant de réflexions issues de la philosophie de la technique. Nous donnerons ensuite un aperçu de ce que sont les outils d'intelligence artificielle et des applications qu'ils peuvent proposer dans le domaine de la santé. Enfin, nous nous tournerons vers une illustration tirée d'un projet de recherche-action actuellement en cours de réalisation et qui s'intitule « Nouvel Humanisme au temps des Neurosciences et de l'Intelligence Artificielle » (NHNAI). Il s'agira de montrer que le « oui » et le « non » à l'égard des IA coexistent effectivement dans nos représentations et que leur mise en tension mérite toute notre attention.

La technicisation de l'humanité.

Les applications informatiques, et parmi elles les outils d'intelligence artificielle (IA), mettent l'humanité face à un choix difficile. Progrès ou danger ? Telle est la question que nous *devons* nous poser, et qui ne cesse en effet de s'inviter un peu partout dans les lieux d'apprentissage et de pratiques professionnelles, mais aussi dans les espaces public et privé.

En Europe en tout cas, la nécessité d'une régulation face aux avancées de l'IA est largement reconnue. Mais pour quelles raisons ? Est-ce la sauvegarde de l'humanité qui est en jeu ? Serait-ce plus largement encore la préservation de l'habitabilité sur notre planète, pour nous *et* les autres vivants ? A moins que ce ne soit pour des raisons plus terre à terre comme l'illustrent les critiques visant la partialité des agents conversationnels, les atteintes à la vie privée et aux droits de propriété intellectuelle. D'un point de vue éthique, ces différents registres de préoccupations ne sont pas équivalents. Selon que l'on adopte le point de vue de certains individus et de leurs droits, ou de la communauté entière des individus humains, ou encore des relations d'interdépendances entre les vivants et leurs environnements, divers devoirs et vertus s'imposeront, se déclinant en autant de manières de dire oui ou non à la nouveauté technique.

L'extraction d'énergie atomique a suscité une perplexité similaire au siècle dernier, et continue en réalité à nourrir de profondes hésitations et à alimenter bien des débats. Comme c'est le cas semble-t-il avec la grande majorité des prolongements techniques que l'être humain engendre, ce n'est qu'après leur mise en œuvre, et lorsqu'ils font sentir leurs effets, qu'un questionnement évaluatif surgit à leur sujet. Et si ces effets s'avéraient plus néfastes que bénéfiques pour nous, pour notre environnement, pour la suite de l'aventure de la vie sur Terre ? A quels types de savoirs ou de sagesses nous en remettre pour guider nos choix dans un monde qui nous contraint déjà à vivre avec ces techniques de plus en plus rapidement envahissantes et agissantes ?

Par ailleurs, comment interpréter le discours des informaticiens qui sont les premiers à démystifier les nouvelles technologies qu'ils produisent, mais qui en même temps ne cessent d'en proposer des versions toujours plus incroyablement proches de certaines de nos capacités humaines ? Comment être rassuré lorsque l'un d'entre eux explique que nous sommes « en pratique à des années-lumière de la création d'un cerveau artificiel capable des mêmes prouesses que notre cerveau organique »¹ ? Voici en effet quelles sont les compétences naturelles de notre cerveau qu'il énumère : « apprendre sans cesse, effectuer un nombre incroyable de tâches différentes, s'adapter aux situations nouvelles, imaginer et créer, ressentir des émotions, etc. » ? Excepté le ressenti des émotions, il n'y a guère d'autre capacité pointée par ce spécialiste en 2016 qui ne soit en 2024 hors de portée d'une IA !

En prenant du recul, nous constatons qu'*homo scientificus*, totalement soumis à la rationalité calculatrice et à la démarche analytique, est en quelque sorte happé par un désir irrésistible qui le domine : la preuve ultime de son intelligence ne serait-elle pas en effet qu'il puisse reproduire celle-ci ? Pour atteindre le véritable statut de maître et possesseur de la nature ne faut-il pas qu'il parvienne à se maîtriser lui-même ? Que vaudrait en effet sa compréhension de l'univers, sa connaissance des lois régissant ce dernier, et toutes les techniques émanant de ce savoir

¹ J-C. Heudin, *Comprendre le Deep Learning. Une introduction aux réseaux de neurones*, Science-eBook 2016, p. 138.

immense, s'il ne régnait pas aussi sur ce qui en constitue la source première, à savoir sa cognition ?

Il n'est au fond pas étonnant que les moyens financiers les plus considérables soient à notre époque dédiés à des progrès technologiques totalement autocentrés, c'est-à-dire entièrement accaparés par ce désir : les neurosciences et l'informatique partent à la conquête du cerveau pour l'imiter (intelligences artificielles) et le prolonger (interfaces cerveau-ordinateur), et à la conquête de nouveaux territoires dans notre système solaire pour assurer la perpétuation de cette « intelligence » se maîtrisant et se possédant en quelque sorte elle-même.

De manière intéressante, le même spécialiste qui rappelait le fossé existant entre l'artificiel et le naturel, reconnaît à ce dernier des spécificités auxquelles nous nous intéresserons dans la suite de cet article. Jean-Claude Heudin reprend un raisonnement que nous rencontrons en fait régulièrement : « l'incroyable complexité de notre cerveau, sa relation intime avec le corps et la perception du monde qui nous entoure, le fait qu'il soit le résultat de plusieurs millions d'années de coévolution avec l'environnement, sont autant d'arguments pour s'assurer qu'il s'agit d'un projet encore hors de portée ».

La modestie et l'humilité du scientifique cachent pourtant mal son désir profond et son indéfectible croyance en la puissance de maîtrise de sa discipline : « un projet *encore* hors de portée ». Nous ne pouvons quant à nous nous empêcher de nous interroger : qu'aurait bien à perdre l'humanité si elle renonçait à ce désir de maîtrise totale ? quel autre désir l'animerait si elle prenait au sérieux son *origine relationnelle* reconnue par certains scientifiques eux-mêmes ? Si c'est effectivement la *coévolution* qui engendre la complexité, si c'est à partir de *réseaux d'interdépendances* qu'est naît cet organe – le cerveau – qui nous fascine tant, pourquoi ne nous intéressons-nous pas avant tout aux relations, et à ces expériences relationnelles qui nous façonnent ?

La relation au cœur du domaine de la santé ?

En prêtant attention aux outils d' « intelligence artificielle », nous sommes invités à nous interroger avant tout sur le rapport entre l'humain et ses productions techniques. Or sur le plan relationnel, une différence profonde apparaît immédiatement : en tant qu'organisme vivant, *l'être humain se reproduit sans visée utilitaire*, alors qu'en tant qu'outil, *la technique n'est jamais produite sans visée utilitaire*. La seule finalité utilitaire ne peut donc pas suffire à rendre compte de toutes les relations qu'expérimentent les vivants, dont font partie les vivants humains. Insistons sur ce point en prenant comme exemple l'expérience propre à la relation de soin.

Écoutons pour ce faire cette parole d'expert : « la médecine est loin de se résumer à la technique, car un patient n'est pas qu'un ensemble d'organes montés dans un corps bien compartimenté ». Ces mots sont ceux de Gérard Berry, tirés de son livre *L'Hyperpuissance de l'informatique* (Odile Jacob, 2017). Cet ingénieur et mathématicien né en 1948, titulaire depuis septembre 2012 de la chaire « Algorithmes, machines et langages » au Collège de France, explicite son point de vue non pas en se servant de ses connaissances scientifiques, mais bien en s'appuyant sur son expérience personnelle :

« J'ai assez d'expérience personnelle comme patient pour savoir qu'un bon médecin a besoin de me voir, d'interagir avec moi, de comprendre comment les différentes parties de mon corps et de mon esprit résonnent entre elles, et pourquoi le soin de la seule partie visiblement en cause ce jour-là peut ne pas suffire. » (p. 218)

Pour bien insister, il clôt le chapitre consacré à « L'informatisation de la médecine » en confiant : « j'espère que je ne serai pas convoqué à l'avenir par un robot chirurgien qui m'opérera sans même s'intéresser à mon mode de vie ! » (p. 219). Et il conclut par une mise en garde adressée à celles et ceux qui prônent de plus en plus une « médecine personnalisée » : celle-ci « ne peut pas être que technique, et il ne faut jamais oublier que personnalisée dérive de personne » (*ibid.*).

Bien que la médecine sorte du champ de son expertise scientifique, Gérard Berry peut comme nous se fier à son vécu d'individu et de patient. C'est ainsi qu'il parvient à toucher un des points les plus sensibles de la technicisation du monde de la santé, à savoir la numérisation des pratiques médicales. Qu'il nous suffise ici de donner comme exemple la présentation de la médecine personnalisée qu'offre l'entreprise pharmaceutique Roche sur son site internet:

« La médecine personnalisée vise aussi à mieux suivre chaque patient, en collectant, mesurant et analysant ses données de santé, pour adapter son cycle de soins à ses besoins. Cette approche innovante repose sur la révolution numérique et notamment le big data. Appliqué à la santé, il permet de collecter les données individuelles avant de les interconnecter, les comparer et réaliser des analyses basées sur l'intelligence artificielle. »²

Avant que nous ne donnions quelques indications sur le fonctionnement de cette fameuse « intelligence artificielle », et de manière à tenter une prise de recul suffisante vis-à-vis de ce qui est brandi comme étant une « révolution numérique », posons-nous à nouveau quelques questions. *Que devient la relation médicale lorsque le patient est approché par le biais d'interconnexions entre des masses de données extraites de lui-même et des autres patients ?*

Essayons de préciser notre interrogation : *la rencontre entre un acteur de la santé et un patient peut-elle encore avoir lieu si le premier substitue à ce dernier son image numérisée, projetée dans une représentation entièrement (in)formée de données et de calculs ? Quelle place donner aux ressentis et aux récits des patients, requérant une expérience relationnelle faite d'écoute et d'empathie, si les analyses produites par les outils d'intelligence artificielle apportent déjà la plupart des réponses permettant de poser un diagnostic ? Que resterait-il de la relation humaine dans un soin qui deviendrait lui-même une pure technique incapable de différencier un patient d'un autre patient autrement que par le biais d'algorithmes numériques ?*

Le message qu'entendait livrer Gérard Berry était le suivant : « Ne confondons pas l'outil et la finalité » (*ibid.*). Or, nous l'avons déjà rappelé, la finalité de l'outil ne peut être que l'utilité. La technique ne peut ainsi se comprendre autrement qu'en étant au service d'autre chose qu'elle-même. Imaginons un médecin qui se mettrait entièrement au service de la technique, par exemple d'une IA exigeant de lui qu'il l'abreuve en données de plus en plus nombreuses. Ce

² <https://www.roche.fr/groupe-roche/vision/medecine-personnalisee> (consulté le 16 mars 2024).

médecin serait alors *lui-même* indéniablement réduit au statut d'outil. Rien ne pourrait alors plus empêcher tôt ou tard son remplacement par d'autres outils plus performants.

Afin que la technique ne soit pas confondue avec la finalité, elle doit être au service d'un but qui est hors de portée pour elle. Ce but ne peut être donné que par les acteurs de la relation de soin eux-mêmes. Il revient bien au système collaboratif médecin-patient de spécifier la finalité de leur rencontre, et de choisir ensemble les moyens (pourquoi pas certaines IA) qui leur permettront de tendre vers ce but.

Pour aborder sereinement la technique

Avant d'explorer plus en détails les techniques d'intelligences artificielles et d'émettre certaines recommandations à leur égard, tournons-nous vers des pensées porteuses d'une sagesse qui aurait tendance à être éclipsée par les paroles savantes de notre époque.

Nous puiserons à deux sources philosophiques bien distinctes et pourtant nettement convergentes. La première raisonnait déjà durant l'antiquité chinoise, au IV^{ème} siècle avant notre ère chrétienne. Nous pouvons l'entendre à travers la traduction d'un récit rapporté par le penseur Tchouang Tseu (Zhuāng Zhōu) :

« Tseu-kong, après s'être rendu dans la principauté de Tch'ou, revenait vers celle de Tsin. En passant au sud du fleuve Han, il vit un vieil homme occupé à travailler son potager. Cet homme descendait par un tunnel dans le puits, en sortait avec sa jarre remplie d'eau et la vidait dans les rigoles de ses plates-bandes. Labeur pénible et mince résultat. Tseu-kong lui dit :

”Si vous aviez une machine avec laquelle cent plates-bandes pourraient être arrosées en un jour, n'aimeriez-vous pas vous en servir ?

- Comment est-ce fait ? demanda le jardinier en levant les yeux sur Tseu-kong.
- Une machine en bois creusé dont l'arrière est lourd et l'avant léger, avec laquelle on lève l'eau comme si on la tirait à la mains, aussi vite que le bouillon déborde de la marmite : cette machine s'appelle une bascule.”

Le jardinier se mit en colère, changea de couleur, ricana et dit : ”J'ai appris de mon maître ceci : qui se sert de machines use de mécaniques et son esprit se mécanise. Qui a l'esprit mécanisé ne possède plus la pureté de l'innocent et perd ainsi la paix et l'âme. Le Tao ne soutient pas celui qui a perdu la paix de l'âme. Ce n'est pas que je ne connaisse pas les avantages de cette machine, mais j'aurais honte de m'en servir.”

Confus, Tseu-kong baissa la tête et ne répliqua pas. »³

Pour celui qui *pense*, le plus grand danger consisterait donc à laisser sa pensée (« son esprit ») se mécaniser par une servitude trop importante à l'égard des machines.

Ce conte simple nous invite à nous questionner à nouveau. Quelle nécessité y aurait-il à faire faire à une machine ce qui occupe ce vieil homme possédant déjà le bien le plus précieux : « la

³ Tchouang-Tseu, L'Œuvre complète, XII « Ciel et terre », Unesco et Editions Gallimard, 1980, p. 170.

paix de l'âme » ? C'est justement en travaillant par lui-même son potager, c'est-à-dire en vivant directement cette expérience relationnelle consistant à prendre soin de la terre, à y confier des semences et à prendre à nouveau soin du fruit de cette rencontre durant sa croissance, que l'homme trouve la sérénité. Pourquoi laisserait-il la technique, en l'occurrence une habile machine, le priver de son activité vitale ?

Il aurait honte de s'en servir, confie-t-il à son interlocuteur. 24 siècles plus tard, nous n'avons pourtant pas honte de nous servir d'innombrables machines et techniques nous remplaçant dans bien des tâches. Nous pouvons même affirmer que ces prouesses technologiques font notre fierté et que nous misons sur elles pour nous accompagner dans notre futur de plus en plus incertain. Que retenir dès lors de cette histoire ? Devons-nous nous sentir coupables ?

Le célèbre physicien Werner Heisenberg, qui commente lui aussi ce conte dans son ouvrage intitulé *La nature dans la physique contemporaine*⁴, accorde son attention à l'incertitude qui frappe l'esprit humain lorsque ce dernier a perdu son innocence et sa simplicité :

« Chacun sent que cette vieille histoire contient une part considérable de vérité ; car "l'incertitude dans les mouvements de l'esprit" est peut-être ce qui décrit de la façon la plus frappante l'état des hommes dans la crise actuelle : la technique, la machine ont envahi le monde dans une mesure dont ne pouvait se douter le sage chinois. » (*ibid.*)

N'accusons toutefois pas trop vite ces propos antiques ou plus récents de technophobie. Heisenberg ajoute que « ce n'est pas à la technique en elle-même qu'on peut imputer la perte du sens de la cohésion de l'ensemble ». De même, du côté du sage chinois : c'est celui qui a laissé son esprit se mécaniser qui perd la paix de l'âme. Étant donné l'état actuel de notre monde et la mécanisation poussée qui le caractérise, nous devrions nous demander non pas comment faire pour diminuer ou supprimer la technique, mais plutôt *comment lui faire une place dans nos expériences relationnelles sans mécaniser nos esprits et nos mentalités*.

Autrement dit, ce n'est pas la technique en elle-même qui est responsable d'une perte de sens, mais bien *l'être humain qui se laisse progressivement accaparer et aveugler par elle*. Comme le reconnaît Heisenberg en y insistant, « pour la première fois au cours de l'histoire, l'homme se trouve seul avec lui-même sur cette terre, sans partenaire, ni adversaire » (*ibid.* pp. 27-28). Ne nous y trompons pas : c'est lorsque l'humain remplace des expériences relationnelles fondamentales (et fondatrices) par des techniques que celles-ci provoquent son isolement dans un monde où tout est réduit à l'utilitaire.

Poursuivons notre réflexion en nous tournant vers un second penseur. Il s'agit du philosophe allemand Martin Heidegger avec lequel Werner Heisenberg fut d'ailleurs en contact. Cette figure incontournable du paysage philosophique récent a proposé une réflexion profonde sur les rapports entre l'humain et la technique. Nous partirons ici d'un discours qu'il a prononcé à Messkirch, sa ville natale, le 30 octobre 1955 et qui fut publié en 1959 sous le titre « Sérénité » (« *Gelassenheit* ») dans le recueil de textes *Questions III*. L'inquiétude à l'égard de la technique y est formulée avec précision :

⁴ W. Heisenberg, *La nature dans la physique contemporaine*, Editions Gallimard, 1962 pour la traduction française, p. 26.

« Ce qui, toutefois, est ici proprement inquiétant n'est pas que le monde se technicise complètement. Il est beaucoup plus inquiétant que l'homme ne soit pas préparé à cette transformation, que nous n'arrivions pas encore à nous expliquer valablement, par les moyens de la pensée méditante, avec ce qui, proprement, à notre époque, émerge à nos yeux. »⁵

A nouveau, ce n'est pas la technique elle-même qui constitue une menace, mais notre rapport à elle : plus précisément le fait d'être préparé ou non à la transformation qu'induit la technicisation du monde. Or cette préparation ne peut pas se faire avec les mêmes moyens qui permettent de produire la technique, à savoir la *pensée calculante* ! C'est pourquoi Heidegger en appelle à un type de pensée particulier, qu'il nomme « pensée méditante » (*besinnliche Denken*), qui est en réalité pour lui la seule et véritable pensée humaine.

Pour nous aider à comprendre en quoi consiste cette pensée contemplative qu'il oppose à la « pensée calculante » (*rechnenden Denken*), Heidegger nous livre l'éclairage suivant :

« La pensée méditante exige de nous que nous ne nous fixions pas sur un seul aspect des choses, que nous ne soyons pas prisonniers d'une représentation, que nous ne nous lancions pas sur une voie unique dans une seule direction. La pensée méditante exige de nous que nous acceptions de nous arrêter sur des choses qui à première vue paraissent inconciliables. »

Retenons pour la suite de notre propos ce caractère *plurivoque* de la pensée méditante, à savoir le fait qu'elle ne se laisse jamais réduire à une seule représentation, à un seul contenu (par exemple le plus savant et le plus complet, ou le plus performant), mais qu'elle reste ouverte aux rencontres par lesquelles adviennent diverses perspectives, quitte à ce que ces dernières paraissent inconciliables.

Selon Heidegger, le véritable danger serait que notre attachement aux techniques soit si fort que nous devenions leurs esclaves. Et le philosophe nous propose de nous y prendre dès lors autrement vis-à-vis des techniques :

« Nous pouvons utiliser les choses techniques, nous en servir normalement, mais en même temps nous en libérer, de sorte qu'à tout moment nous conservions nos distances à leur égard. Nous pouvons faire usage des objets techniques comme il faut qu'on en use. Mais nous pouvons en même temps les laisser à eux-mêmes comme ne nous atteignant pas dans ce que nous avons de plus intime et de plus propre. »

La voie prônée par Heidegger, loin du technophobie radical dont il est parfois affublé, consiste à *dire à la fois oui et non aux techniques* : *oui* à leurs apports indéniables, à leur importante participation dans certaines de nos activités, et en même temps *non* afin de ne pas les laisser envahir notre esprit, nous priver d'expériences relationnelles essentielles, et ainsi « fausser, brouiller et finalement vider notre être ».

Cette voie est celle que Heidegger nomme « sérénité », et qui renvoie selon lui à « l'égalité d'âme ». Sans entrer dans le débat au sujet de la traduction du vieux terme allemand « *Gelassenheit* », soulignons l'importance du geste ainsi esquissé par le philosophe. Loin de

⁵ Éditions Gallimard 1966 pour la traduction française. Traduit de l'allemand par André Préau.

rendre ambigu notre rapport aux choses techniques, cette réponse de « oui et non » consiste tout simplement à les admettre dans notre environnement quotidien tout en les laissant dehors, « comme des choses qui n'ont rien d'absolu, mais qui dépendent de plus haut qu'elles ».

Bref, il s'agirait de rester maîtres et possesseurs de nos techniques, sans nous laisser intimider par leurs prouesses imitant en fin de compte notre seule pensée « calculante », et non pas la pensée « méditante ».

La sérénité, cette égalité d'âme dont parle Heidegger, n'est-elle pas proche de la paix de l'âme qu'évoquait le sage chinois ? Aussi mystérieuse qu'elle puisse encore nous paraître, n'est-ce pas d'une pensée méditante dont nous avons le plus besoin aujourd'hui pour envisager sereinement notre rapport à la technique, c'est-à-dire pour faire face à l'incertitude dans laquelle elle nous plonge, et cela en nous rendant capables de lui dire à la fois oui et non ?

Pointons encore un dernier écho à cette réflexion dans les propos récents de Raphaël Enthoven pour qui « plus nous nous livrons à la tutelle des écrans, plus nous hallucinons leur propre autonomie ». Et le philosophe français de poursuivre dans la même interview :

« Le rêve que nos machines sont vivantes accompagne la machinisation de nos esprits. Comme si nous n'étions pas seuls responsables d'une telle hypnose ! »⁶

A nouveau, c'est cette machinisation de l'âme par la technique dont nous sommes seuls responsables qui est ici mise en évidence. Et nous retrouvons par ailleurs dans les propos du même philosophe l'importance de la dimension relationnelle humaine :

« De façon générale, l'IA ne saurait synthétiser une relation. Car aucune relation, même la plus intime, ne permet de prédire à la virgule près le contenu d'une discussion. » (ibid.)

Qu'est-ce que l'IA ?

Les techniques d'intelligence artificielle qui nous impressionnent tant, notamment par leurs capacités génératives de plus en plus avérées, ont une histoire particulière qu'il s'agit de retracer brièvement pour bien comprendre ce qui est au cœur de leur fonctionnement.

Il nous suffira de faire commencer cette histoire avec la cybernétique née au milieu du siècle dernier. Norbert Wiener (1894-1964) avait alors mis au point des méthodes s'appliquant à une vaste classe de problèmes relevant aussi bien de la télécommunication, du comportement des servomécanismes et des ordinateurs, que du système nerveux. Il proposa alors le terme *cybernétique* pour caractériser la théorie naissante de « la communication et de la commande chez l'animal et dans la machine ». Dès la publication en 1948 de son livre *Cybernetics*, ses idées connurent un succès considérable et trouvent encore aujourd'hui des prolongements dans le domaine des sciences cognitives.

Si ce courant perd progressivement son nom, il n'en va pas de même pour *l'approche connexionniste* sur laquelle il fonde ses espoirs de reproduire artificiellement les compétences cognitives du système nerveux humain : les réseaux de neurones artificiels font en effet leur

⁶ <https://www.philomag.com/articles/raphael-enthoven-lesprit-est-cette-etrange-etoffe-dont-lhumanite-est-faite-et-que-la>

apparition et le perceptron est inventé en 1957 par Frank Rosenblatt. Le paradigme connexionniste, qui utilise la physique statistique et s'inspire d'une partie du fonctionnement du cerveau, propose des unités réalisant des calculs simples, mais que l'on peut démultiplier en parallèle et enchaîner les unes aux autres. Un neurone artificiel fait ainsi une sorte de calcul test : en entrée, il prend une série de nombres qui sont multipliés chacun par un coefficient réglable avant d'être ajoutés. Si le résultat de cette somme est en dessous d'une valeur seuil elle-même réglable, il ne se passe rien. Si la valeur seuil est atteinte ou dépassée, le calcul renvoie un nombre non nul en sortie (souvent au voisinage de 1).

Ce point de départ est très important car lorsque John McCarthy et Marvin Minsky proposent le terme d'« intelligence artificielle » en 1956 lors de la réunion fondatrice de Dartmouth, c'est précisément pour l'opposer au connexionnisme de la cybernétique. Pour eux, « tout aspect de l'apprentissage ou toute autre caractéristique de l'intelligence peut en principe être décrit avec une telle précision qu'une machine peut être fabriquée pour le simuler »⁷. Leur ambition était alors de « trouver comment faire en sorte que les machines utilisent le langage, forment des abstractions et des concepts, résolvent des types de problèmes aujourd'hui réservés aux humains et s'améliorent elles-mêmes. » (*ibid.*). Leur approche était résolument *symbolique*.

L'idée des fondateurs de cette IA « symbolique », parfois connue également sous l'appellation d'« IA forte », mais considérée aujourd'hui comme étant l'« IA à l'ancienne », est de mettre dans les programmes des ordinateurs des règles leur permettant de manipuler des représentations de haut niveau.

Toutefois, au début des années 1970, l'IA connut un « premier hiver » qui gela la plupart des projets tant symboliques que connexionnistes : « Les deux courants ont beaucoup trop promis et les résultats sont loin d'être au rendez-vous. »⁸.

Elle renaîtra dans les années 1980 après avoir révisé de manière profonde l'architecture des machines symboliques auxquelles elle donnera le nom de « systèmes experts ». De telles IA s'appliquaient alors essentiellement à des contextes scientifiques et industriels. *Mycin* est le premier et le plus connu de ces systèmes. Il fut conçu à l'université de Stanford et s'appuie sur une base de connaissances de 600 règles permettant de traduire le raisonnement humain aboutissant au diagnostic de maladies infectieuses du sang.

A partir du début des années 1990, le paradigme connexionniste fait son retour suite à un déplacement qui s'opère au sein du champ de l'apprentissage artificiel. Ce dernier, alors en plein essor, commençait à donner naissance à des algorithmes de classification de plus en plus performants réunis sous l'appellation *Machine Learning*. Mais un changement dans la nature des données composant le domaine d'entrée de ces machines explique le retour du débat entre approches symbolique et connexionniste.

Non seulement plus nombreuses, ces nouvelles données n'étaient en outre plus porteuses de structure intrinsèque. Elles ne pouvaient dès lors plus être rassemblées que sous la forme d'un

⁷ McCarthy, J., Minsky, M., L., Rochester, N., Shannon, C., E., « A pro-posal for the Dartmouth summer research project on artificial intelligence » (1955), réédité dans *AI Magazine* 27/4 (2006).

⁸ Cardon, D., Cointet, J. & Mazières, A. (2018). La revanche des neurones : L'invention des machines inductives et la controverse de l'intelligence artificielle. *Réseaux*, 211, 173-220. <https://doi.org/10.3917/res.211.0173>

voisinage statistique. De telles données ne donnant plus d'informations sur les relations qu'elles entretiennent entre elles perdaient leur caractère « symbolique » :

« L'horizon du calcul n'est plus interne au calculateur, mais est une valeur que le monde lui donne de l'extérieur – et qui est souvent très « humaine », comme va en témoigner tout ce travail manuel d'étiquetage des données : cette image contient-elle (ou pas) un rhinocéros ? Cet utilisateur a-t-il cliqué (ou pas) sur tel lien ? La réponse (le critère d'optimisation) doit être entrée dans le calculateur avec les données pour que celui-ci découvre un « modèle » adéquat. » (*ibid.*)

Les informaticiens néo-connexionnistes à l'origine de ce basculement imposent également leurs propres termes dans le débat avec les tenants de l'IA symbolique. Ils opèrent ainsi une distinction entre la « largeur » des architectures « minces » propres aux systèmes d'IA symboliques (programmation logique, logique de description, etc.), et la « profondeur » des architectures en couches de neurones qu'ils proposent. Le terme de *Deep Learning* est ainsi forgé en 2006 par Geoff Hinton.

Alors qu'avec le *Machine Learning* classique, le Data Scientist doit lui-même faire un choix, et extraire la donnée qui va influencer sur la prédiction opérée par l'algorithme, ce ne sera plus du tout le cas avec le *Deep Learning*. Ici, plus besoin de choisir, selon les variables souhaitées, les données avec lesquelles l'algorithme va être entraîné, car les données vont pouvoir être apportée à l'algorithme de manière brute.

Le « vrai monde », celui des données produites par la numérisation des images, des sons, des paroles et des textes, est non linéaire : il contient du « bruit », l'information y est redondante, les flux de données ne sont pas catégorisés derrière des attributs de variables homogènes, claires et construites de façon intelligible, les exemples sont parfois faux, etc. Or c'est à ce monde « réel » là que les connexionnistes veulent confronter leurs réseaux de plus en plus gigantesques de neurones artificiels afin qu'ils puissent « apprendre à identifier et à démêler les facteurs explicatifs sous-jacents cachés dans le milieu observé des données sensorielles de bas niveau »⁹.

Dans une interview accordée à l'occasion de la parution de son livre *Parole de machines* (humensciences, 2023), le chercheur et spécialiste Alexei Grinbaum résume ainsi les résultats étonnants que cette approche connexionniste permet d'atteindre :

« Ce qui est intéressant dans cette science, c'est qu'on ne sait pas pourquoi cela marche. On augmentait la taille des réseaux de neurones et, à un moment donné, les sorties sont devenues quasi parfaites. Mais pourquoi ? On est certain que ce phénomène relève de la physique statistique. (...) Le système GPT-2 avait trois milliards de paramètres et GPT-3 atteignait 175 milliards de paramètres : entre les deux, le saut dans la qualité des résultats a été merveilleux. »¹⁰

Au fond, ce que nous apprennent ces IA converties au connexionnisme relevant du *Deep Learning*, c'est que nos univers langagiers et sensoriels – qui comprennent nos modes de

⁹ Yoshua Bengio et al. (2013) ...

¹⁰ « Parler avec les machines. Entretien avec Alexei GRINBAUM » : <https://www.revue-etudes.com/article/parler-avec-les-machines/26324>

communication introspectif (pensée) et intersubjectif (dialogue) – sont de nature profondément relationnelle. Le même auteur nous éclaire davantage sur ce point :

« L'acte de donner un nom consiste à établir une relation et à mettre l'autre dans son monde. Peu importe qu'il soit un animal, un oiseau, une machine ou je ne sais qui : ce n'est pas le statut ontologique qui compte, *c'est la relation*. (...) L'exemple le plus parlant, ce sont les prêtres du sanctuaire de Delphes. Ils interprètent la parole asémantique de la Pythie. Le prêtre dit, par exemple : *"Je m'appelle Plutarque et je vais vous dire quel est le jugement d'Apollon"*. C'est la même histoire pour la machine : son résultat est asémantique et un être humain l'interprète et lui donne sens. » (*ibid.*)

Du côté de la santé, le problème de l'exécution par des machines de certaines tâches médicales répétitives, et donc sujettes à l'erreur humaine, était en bonne voie d'être résolu dès les années 1990 et au début des années 2000. Selon les médecins Charlotte J. Haug et Jeffrey M. Drazen, la lecture informatique des électrocardiogrammes (ECG), le comptage différentiel des globules blancs, l'analyse automatisée de photographies rétinienne et de lésions cutanées, ainsi que d'autres tâches de traitement d'images, sont donc déjà des réalités, et

« [p]our ces applications, il a été démontré que l'IA et l'apprentissage automatique aident le prestataire de soins de santé en signalant les aspects des images qui s'écartent de la norme. »¹¹

Mais pour ces auteurs, de tels progrès suscitent un questionnement crucial :

« Cela soulève une question essentielle: qu'est-ce que la norme ? Cette simple question montre l'une des faiblesses de l'utilisation de l'IA et de l'apprentissage automatique en médecine telle qu'elle est largement appliquée aujourd'hui. » (*ibid.*)

Cette question de la norme pourrait donc bien nous ramener à notre point de départ : dire à la fois oui et non à l'utilisation des IA. Étant donné que le « non » et le « oui » s'éclairent l'un l'autre, voyons ce que cette question peut nous apprendre à leur sujet.

Qu'est-ce que la norme ?

Servons-nous de cette question pour poursuivre notre exploration d'une manière de nous rapporter à la technique, et à l'IA en particulier, qui consiste à lui dire à la fois oui et non.

Rappelons qu'à côté du « oui » qui est déjà là – nous venons encore de le confirmer dans la section précédent en ce qui concerne le domaine de la médecine – nous cherchons à établir la nécessité d'un « non » qui ne vienne pas s'y opposer et le neutraliser, mais bien plutôt le compléter et nous permettre de retrouver une dimension essentielle de notre condition d'être humain, et même d'être vivant. L'enjeu est de parer le risque de machinisation de notre esprit (dire uniquement « oui » à la technique) en y maintenant une place pour la pensée méditante ouverte à la pluralisation des perspectives, des représentations, des choix (pouvoir toujours dire « non », malgré que nous disions « oui »).

¹¹ Haug CJ, Drazen JM. Artificial Intelligence and Machine Learning in Clinical Medicine, 2023. N Engl J Med. 2023 Mar 30;388(13):1201-1208. doi: 10.1056/NEJMra2302038. PMID: 36988595. Notre traduction.

Aussi puissantes soient-elles pour produire ce qui ressemble à notre langage et à notre monde sensoriel, les techniques d'IA restent totalement incapables d'y reconnaître du sens. Elles sont limitées à la pensée *calculante* et ne pensent dès lors pas, pour reprendre les termes de Heidegger. L'informaticienne Isabelle Linden, professeure à l'Université de Namur, résume de la manière suivante cette limite inhérente à la simulation informatique :

« Aujourd'hui, bien des facettes de l'être humain échappent encore à la modélisation formelle : la conscience de soi, les émotions, les sentiments, les relations sociales... Pour les intégrer dans une IA, il serait nécessaire de les exprimer dans des termes matériels et il faudrait que ces formulations soient calculables. »¹²

D'une certaine manière, c'est en cherchant à se dédoubler dans des techniques comme celles de l'IA que l'être humain peut le mieux prendre conscience de qui il est réellement. C'est cette piste que nous poursuivrons par la suite, et qui peut déjà dans la présente section nous aider à approcher la capacité normative proprement humaine.

Dans une tribune parue récemment dans le *New York Times*¹³, le philosophe et linguiste Noam Chomsky, père de la grammaire générative, donne un petit exemple grammatical permettant de comprendre la servilité totale d'une IA la rendant incapable d'accéder à quelque norme que ce soit. Nous reprenons la version de cet exemple présentée par Martin Legros dans *Philosophie Magazine*¹⁴.

Tout anglophone qui entend la phrase « *John is too stubborn to talk to.* » comprendra immédiatement le sens de cette phrase grâce à sa maîtrise de la langue et à la situation dans laquelle elle est énoncée. Elle voudra dire : « *John est trop têtu pour qu'on le raisonne.* ». Dans cette compréhension, « John » joue un double rôle grammatical : il est sujet de la phrase, mais devient également complément d'objet direct. Par ailleurs, cette compréhension nous autorise à traduire « talk » par « raisonner » et non « parler ». Or une IA comme ChatGPT traduira cette phrase par « *John est trop têtu pour parler à quelqu'un.* ». N'ayant accès ni aux règles, ni aux contextes d'énonciation, son calcul statistique prédira un énoncé « optimal » uniquement déterminé à partir d'une quantité énorme d'occurrences trouvées dans de vastes corpus et présentant une proximité analogique (selon une démarche purement inductive).

Chomsky utilise cet exemple pour montrer que « ChatGPT fait preuve de quelque chose comme la banalité du mal : plagiat, apathie, évitement [...]. Ce système offre une défense du type "je ne fais que suivre les ordres", en rejetant la responsabilité sur ses créateurs. » (*ibid.*). Et Martin Legros conclut quant à lui ceci :

« Notre intelligence ne se contente pas de définir ce qui est ou ce qui pourrait être ; elle cherche à établir ce qui doit être. » (*ibid.*)

Nous proposons donc de placer cette compétence humaine normative (évaluative) du côté de la pensée non calculatrice, « méditante », à savoir celle qui ne se fixe jamais sur *un seul* point de vue, qui ne se laisse pas enfermer dans *une* représentation (la plus probable, la plus optimale,

¹² Linden, I., Tilman, V. & Laurent, N. (2023). Les techniques d'intelligence artificielle : histoire, développements et défis. *Recherches de Science Religieuse*, 111, 603-624. <https://doi.org/10.3917/rsr.234.0603>

¹³ <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>

¹⁴ <https://www.philomag.com/articles/chatgpt-chomsky-et-la-banalite-du-mal>

autrement dit la plus convexe en langage d'informaticien), et qui s'accommode de ce qui en apparence paraît inconciliable. Autrement dit, si en tant qu'êtres humains nous ne pouvons jamais manquer de donner sens à ce que nous rencontrons (nous *interprétons* continuellement), c'est parce que *plusieurs* sens se présentent toujours à nous en même temps. C'est donc dans chaque rencontre que surgit la question de la norme : quel sens dois-je donner à cela ? quelle valeur lui attribuer ? que dois-je en comprendre ?

Pour approfondir cette réflexion au sujet de la norme, tournons-nous à présent vers le philosophe français de la médecine Georges Canguilhem (1904-1995). Celui-ci a en effet critiqué l'idée selon laquelle la différence entre le normal et le pathologique serait réductible à une simple variation quantitative de certains paramètres physiologiques. Une telle idée est présumée par les approches statistiques en médecine qui envisagent le normal comme un état physiologique moyen et le pathologique comme un écart par rapport à ce dernier. Bien entendu les outils d'IA s'inscrivent parfaitement dans ce courant « calculateur » en venant servir cette approche normative généralisatrice.

Or pour Canguilhem les êtres vivants doivent être considérées « moins comme des êtres référables à un type réel préétabli, que comme des organisations dont la validité, c'est-à-dire la valeur, est référée à leur réussite de vie éventuelle »¹⁵. Pour lui, « normal » ne désigne pas un aspect de l'existence susceptible d'être mesuré et déterminé objectivement. Une norme n'est pas nécessairement existante en tant que telle, et dès lors l'état de santé qui y est associé ne peut pas exister non plus en tant que tel. Une norme ne découle pas d'un jugement sur la réalité, mais plutôt d'un jugement concernant une certaine valeur. Elle met en rapport ce qui existe avec ce qui doit exister.

Canguilhem renverse donc la conception traditionnelle de la norme. Il rejette la thèse selon laquelle la vie serait produite par des normes agissant sur elle de l'extérieur, en affirmant que les normes sont elles-mêmes produites par le mouvement de la vie. Les normes se constituent dans le cadre d'un ordre engendré par l'organisme lui-même.

Selon ce philosophe, tant les normes que les valeurs émergent de la rencontre entre les exigences de l'organisme et les conditions environnementales. La normalité n'est pas une adaptation à des normes préexistantes, mais plutôt une capacité d'instaurer de nouvelles normes au fil des rencontres entre les valeurs de l'organisme et les conditions environnementales, ou entre les valeurs de plusieurs organismes.

Pour Canguilhem, la santé est donc bien plus qu'un état physiologique normal. Elle implique en effet une *capacité normative*, c'est-à-dire une capacité de créer de nouvelles normes. La santé renvoie ainsi à l'aptitude de tolérer des fluctuations, de surmonter des contradictions, et finalement de donner du sens¹⁶. Un organisme dit « malade » est donc enfermé depuis une perspective extérieure (objectivante) dans des conditions d'existence définies et fixes, en faisant comme s'il avait perdu sa capacité normative.

¹⁵ G. Canguilhem. « Le normal et le pathologie », in *La connaissance de la vie*, Librairie philosophie J. Vrin, 1965.

¹⁶ Le biologiste Olivier Hamant reprend et développe cette idée en l'associant à la notion de « robustesse ». Voir ses ouvrages récents : *La troisième voie du vivant* et *De l'incohérence* (Odile Jacob).

Aller à la rencontre du « oui » et du « non »

Nous voudrions illustrer la coexistence du « oui » et du « non » à l'égard de la technique en mentionnant l'apport d'un projet de recherche-action mené conjointement par plusieurs universités à l'échelle internationale, et auquel nous participons depuis un peu plus de deux ans.

Ce projet interuniversitaire et international est dirigé par l'Université Catholique de Lyon, et plus précisément par un des auteurs de cet article, Mathieu Guillermin. Il s'intitule « Nouvel humanisme au temps des neurosciences et de l'intelligence artificielle », et vise à mettre des communautés d'acteurs de la société en capacité de conduire une réflexion collective sur ce que signifie être humain à notre époque marquée par l'arrivée de techniques introduisant de nouveaux modes d'interactions avec nous, parfois très intrusifs (songeons aux implants neuronaux).

Dans cette recherche-action, nous nous intéressons à notre préconception de ce que signifie être humain, à savoir un ensemble de « préjugés » (nous préciserons plus loin ce terme) reflétant ce que nous pouvons appeler l' « humanisme » et qui sont à l'œuvre implicitement ou explicitement lorsque nous agissons et fixons un cap ou des limites. Cet « humanisme » ne peut manquer d'être profondément mis en tension par ces avancées en neurosciences et en IA. Il apparaît donc crucial que ce pilier participant à la réflexion éthique induite par le développement rapide des champs de l'IA et des neurosciences soit davantage rendu explicite.

Le cœur du projet NHNAI consiste à animer et soutenir une réflexion *avec* les acteurs sociétaux sur cette question de l'humanisme au temps de l'IA et des neurosciences. Trois champs d'application ont été choisis pour faire l'objet d'une attention toute particulière, à savoir l'éducation, la santé et la démocratie.

Un tel projet ne peut faire l'économie d'une réflexion approfondie au sujet des *méthodes* envisagées pour mettre en œuvre une réflexion collective : comment faire vivre cette dynamique d'échanges qui, bien que supposée présente au sein de toute démocratie, n'en reste pas moins une expérience difficile et rare nécessitant :

- La création d'*espaces publics de rencontres* (espaces évidemment spatio-temporels).
- Des échanges de paroles fondés sur l'écoute et la compréhension mutuelles, et donc *l'accueil de points de vue forcément situés* reflétant toutes sortes d'expériences vécues entremêlées.
- La mise en évidence de convergences, mais aussi et surtout de tensions qui font apparaître *les points de vue en tension*.

Pour rappel, ce sont ces derniers qui sont caractéristiques d'une pensée « méditante » capable de dire oui et non en même temps à la technique. Reconnaître l'existence de points de vue en apparence inconciliables et apprendre, grâce à eux, à réfléchir et à agir ensemble, nous paraît être une boussole éthique importante. C'est en tout cas le guide choisi par ce projet pour tenter de faire apparaître de nouvelles manières de vivre humainement avec l'aide de la technique.

A sa manière, ce projet concrétise une idée que Bernadette Bensaude-Vincent avait avancée dans son ouvrage *Les vertiges de la technoscience* (La Découverte, 2009). Selon la philosophe et historienne des sciences le « vertige de la technoscience » « provient de ce que la

technoscience construit un monde plat, sans relief sans repères, un monde lisse, peuplé d'automates performants et peu pensants » (p. 197). Elle en appelle ainsi à « reconstituer un paysage avec des reliefs pour constituer un monde habitable », et à « recréer un monde sous tension pour résister au processus d'uniformisation » (*ibid.*). Pour elle comme pour nous, pareil objectif ne peut être *uniquement* confié aux experts (spécialistes, scientifiques). Rien ne nous garantit en effet qu'ils puissent nous guider vers autre chose que cette technoscience vertigineuse. Nous nous rallions donc résolument au projet de « civilisation » appelé de ses vœux par l'autrice, étant entendu que « civiliser cela veut dire, en premier lieu, engager la société civile, faire participer les citoyens » (*ibid.*).

Nous ne nous avancerons pas ici dans un travail d'analyse, mais souhaitons simplement présenter une série de points de vue divergents que nous avons recueillis lors de rencontres avec des publics très variés (en âges, genres, occupations professionnelles, formations, etc.).

Dire « non » au nom du relationnel :

« La technologie engendre une modification de la manière de travailler du médecin : on pourrait même dire qu'elle la formate (ex. encodage), et cela peut se faire au détriment du relationnel et du temps passé à communiquer directement avec le patient. »

« L'IA et la numérisation utilisées à l'excès pourraient faire obstacle à certains mécanismes humains tels que les relations de confiance et la responsabilité qui sont des atouts dans le monde professionnel. En effet, ce sont de petits signaux, parfois presque imperceptibles, qui permettent aux humains de prévenir des problèmes, de rectifier des processus. »

« L'appui des technologies, le diagnostic et la prescription sont gérés différemment. Quel rôle donner à la technologie ? Il est essentiel en tout cas de préserver la relation humaine. Il ne faudrait pas laisser de côté ou brider l'expression du patient et la réponse personnelle du médecin. Il est essentiel de tenir compte du relationnel, de la rencontre entre humains, de la communication verbale et non verbale. Pour de multiples raisons, dont celles qui viennent d'être suggérées, la télémédecine, par exemple, ne devrait pas remplacer une vraie rencontre avec un vrai médecin »

« Je crains que la dimension relationnelle soit mise de côté si le médecin reçoit une réponse (diagnostic) à partir de données encodées dans une machine. L'expérience et l'intuition du médecin seraient alors beaucoup moins utilisées. Autrement dit, le savoir-faire du médecin pourrait être amené à moins se développer. Cette préoccupation devrait avoir pour conséquence, non pas de refuser d'utiliser une technologie susceptible d'apporter des bénéfices au patient, mais de réfléchir aux implications des choix technologiques et d'implémenter les technologies en fonction ou en tenant compte de ces implications. »

Dire « oui » au nom de l'efficacité :

« Une cohabitation intelligente avec les outils technologiques de plus en plus puissants est certainement envisageable aux conditions suivantes : si ces outils permettent de gagner du temps en faveur des relations humaines ; si les risques de dérives ne sont pas ignorés ; si les autorités publiques jouent leur rôle. »

« Les médecins (surtout les spécialistes) ont actuellement peu de temps à consacrer aux patients et les technologies peuvent effectivement leur apporter une aide de ce point de vue. »

« Il faut oser l'évolution si nous en obtenons des bénéfices. Par exemple, les nouvelles technologies peuvent être justifiées par le fait qu'elles font gagner du temps au médecin, si c'est au bénéfice des relations humaines et surtout de la demande du patient. »

« Toutefois, une réaction surgit face à cette affirmation : un programme informatique suffisamment bien adapté ne serait-il pas toujours efficace (plus efficace qu'un humain) ? »

« Il y a également l'opportunité de voir se produire l'effet inverse : à savoir que, grâce à l'IA, le médecin puisse bénéficier de l'expérience cumulée de médecins spécialisés. Par là, sa compétence pourrait au contraire se développer. »

« Dans le cas où un patient a passé de nombreux tests, et que suffisamment de données sont disponibles, alors l'IA pourrait éventuellement lui apporter une aide importante, par exemple pour appréhender plus efficacement les interactions entre différents traitements. »

« La question qui compte est celle-ci : une technologie peut-elle sauver des vies ? Si oui, alors il ne serait pas éthique de refuser de l'utiliser. »

« Globalement, on peut dire que le bilan des technologies est positif : nous sauvons de plus en plus de vies, notamment en luttant de plus en plus efficacement contre les maladies liées au vieillissement. »

Conclusion

Dans cet article nous avons approché le rapport entre les techniques d'Intelligence Artificielle et la relation de soin par le biais d'une réflexion plus large au sujet de la technique.

Plutôt que d'opposer partisans et opposants d'une technologisation toujours plus poussée du domaine de la santé, et pour éviter un débat stérile qui consisterait à mettre en scène un faux dialogue entre arguments se neutralisant les uns les autres, nous avons proposé une autre voie. En partant d'une réflexion philosophique passant notamment par la sagesse Taoïste et la pensée de Martin Heidegger, nous avons tenté d'évaluer dans quelle mesure nous pouvons – et même devons – dire « oui » *et* « non », en même temps, à la technique, et plus particulièrement à celle qui se propage actuellement à une vitesse vertigineuse sous formes d'algorithmes apprenants (Machine et Deep Learning).

Le plus souvent, nous disons « oui » aux techniques et notamment aux IA, parce qu'elles peuvent s'avérer utiles. En cela elle s'inscrivent dans des projets humains à finalités humaines. Ce « oui » implique donc de toujours pouvoir demander aux concepteurs de justifier les moyens mis en œuvre pour atteindre leurs objectifs. La transparence des fins est essentielle pour assurer cette première raison de dire « oui ».

Dans le domaine de la médecine, nous pouvons illustrer ce « oui » en mentionnant une classification d'outils d'aide au diagnostic faisant appel aux IA, élaborée pour favoriser leur émergence¹⁷ :

« L'assistance par l'IA (IA-A), où l'IA apporte des datas complémentaires pour aider l'analyse diagnostique humaine ;

La substitution par l'IA (IA-S), où l'IA remplace le travail diagnostique humain et formule lui-même des hypothèses basées sur la donnée ;

¹⁷ M. Durand, A. Shaikh, M. Billi, E. Lechevallier, Applications médicales de l'intelligence artificielle : opportunités & challenges, Progrès en Urologie - FMC, Volume 30, Issue 2, 2020, <https://doi.org/10.1016/j.fpurol.2020.02.001>.

L'extension par l'IA (IA-X), où l'IA contribue à permettre la réalisation de nouvelles tâches médicales au-delà du champ des possibles actuels. »

Mais une seconde raison de dire « oui » aux techniques d'IA apparaît, qui nous intéresse tout particulièrement. Elles nous aident en effet à mieux nous *comprendre* nous-mêmes en révélant au grand jour (en concrétisant) les *préjugés*¹⁸ qui nous permettent d'entrer en contact avec le monde et autrui. En l'occurrence, avec les techniques de *Machine* et de *Deep Learning*, ce sont les préjugés à l'égard de nous-mêmes qui sont en jeu, et c'est désormais à travers eux que nous entrons en relation les uns avec les autres. Les outils d'IA, en particulier, reflètent avec une grande netteté certains de ces préjugés les plus élaborés à notre égard. Que ce soit via l'approche symbolique ou celle connexionniste, ces outils fonctionnent selon des logiques calculatrices trahissant *une certaine manière de nous représenter* notre « intelligence », à savoir nos capacités cognitives.

Dans le contexte de la relation de soin, ce point mérite toute notre attention. Rappelons-nous la question que nous nous posions : *la rencontre entre un acteur de la santé et un patient peut-elle encore avoir lieu si le premier substitue à ce dernier son image numérisée, projetée dans une réalité entièrement (in)formée de données et de calculs ?* Si un tel préjugé devient nécessaire au regard d'une médecine progressiste, est-il néanmoins suffisant au regard d'une médecine qui se veut toujours humaine ?

L'idée sous-jacente est donc que les artefacts simulant certaines capacités cognitives humaines peuvent être considérés comme le miroir de nos préjugés à l'égard de nous-mêmes. Avec eux nous assistons vraisemblablement à l'expansion d'un véritable *anthropomimétisme*¹⁹. A cet égard, nous rejoignons les idées de Philippe Breton que nous retrouvons dans l'ouvrage paru en 1995 sous le titre *A l'image de l'Homme. Du Golem aux créatures virtuelles* (éditions du Seuil) :

« à travers les créatures, façonnées justement à l'image de l'homme, on peut discerner les multiples représentations de l'humain qui ponctuent, en les structurant, les cultures, essentiellement occidentales. » (p. 6)

De manière intéressante, l'auteur suggère dans la conclusion de son livre que la représentation que l'être humain a de lui-même serait peut-être « en train de se collectiviser » :

« La représentation de l'humain comme sujet individuel laisse ainsi progressivement la place à une représentation de l'homme en tant qu'être social. » (p. 178)

Pouvons-nous dire en 2024 que l'IA est bien devenue cette sorte de « prothèse sociale » (et donc plus uniquement un prolongement de l'individu) entrevue en 1995, à une époque où elle était encore très balbutiante ? Répondre à cette question nous emmènerait bien au-delà des

¹⁸ Nous employons ici ce terme dans le sens positif que lui a donné Hans-Georg Gadamer dans *Vérité et méthode* (1960). Les préjugés sont, selon ce philosophe allemand, des éléments constitutifs du comprendre, c'est-à-dire de notre être-au-monde. Constitutifs de notre manière d'entrer en relation avec le monde et autrui, ils sont la condition de toute compréhension, dans le sens où ils nous ouvrent un accès à ce que nous pouvons comprendre (interpréter).

¹⁹ Compris dans un sens équivalent au biomimétisme : imitation des procédés naturels et de leurs dynamiques pour inspirer de nouvelles innovations techniques.

limites que nous nous sommes fixées dans le cadre de cet article. Revenons-en plutôt à notre approche du « oui et non ».

Nous pouvons relier de la manière suivante les deux raisons de dire « oui » aux techniques d'IA : d'une part nous reconnaissons une utilité à ces simulateurs de tâches cognitives, et d'autre part nous révélons par la même occasion les présupposés que nous avons sur nous-mêmes.

Ce « oui » est important, étant donné que c'est grâce à nos préjugés que nous rencontrons le monde et autrui²⁰. Mais un tel « oui » appelle alors nécessairement un « non », car si les préjugés rendent possible nos rencontres et nous permettent de faire connaissance avec le monde et autrui, ils ne justifient pas pour autant notre attitude dominatrice : nous n'avons nullement le droit de sommer le monde et autrui de se soumettre au diktat de nos préjugés. Notre compréhension de nous-mêmes, en particulier, serait ainsi bien appauvrie si elle se limitait uniquement à ce que nos artefacts parviennent à en reproduire.

Dire « oui » aux techniques d'IA doit donc entraîner un « non » : une fois les préjugés reconnus, nous avons le *devoir* de ne pas les laisser envahir le champ expérientiel de nos rencontres avec le monde et avec autrui. Le caractère *humain* de nos expériences relationnelles rejaillit ici avec la plus vive tonalité, et nécessairement sous forme d'un questionnement : que vais-je apprendre du monde et de l'autre que je ne puis entendre si je limite ma compréhension à mes préjugés ? Plus profondément, et sur un plan éthique et non plus seulement épistémologique cette fois, cette autre question surgit : que dois-je apprendre du monde et de l'autre pour pouvoir continuer à vivre avec eux ?

Cette dernière question résonne tout particulièrement à notre époque où les crises s'enchaînent, et concernent notamment le maintien d'une habitabilité de notre planète terre pour l'être humain et tous ses autres habitants. Que devons-nous apprendre d'eux pour pouvoir continuer à vivre ensemble sans nous laisser enfermer par des préjugés que nous aurions omis de faire évoluer ? Signalons simplement à cet égard qu'il n'est aujourd'hui pas du tout évident qu'un recours massif à des techniques d'IA de plus en plus puissantes (notamment celles basées sur le *Deep Learning*) soit compatible avec le respect de notre environnement. Certains auteurs ont par exemple montré que l'empreinte carbone liée à l'entraînement de modèles algorithmiques profonds utilisés en traitement du langage naturel est cinq fois plus forte que celle d'une voiture durant sa période moyenne d'utilisation²¹. Ce sujet est également abordé dans un article déjà mentionné précédemment²².

A l'échelle de la relation de soin, ces « oui » et « non » à la technique doivent également coexister. L'expérience relationnelle que vivent le soignant et le soigné est en effet l'une des plus propices à nous faire apprendre comment dépasser nos préjugés. Dire parfois « oui », mais aussi parfois « non », à l'aide au diagnostic faisant appel à des IA devient ainsi le devoir qui

²⁰ Nous renvoyons à nouveau sur ce point aux thèses de Hans-Georg Gadamer dans *Vérité et méthode*.

²¹ Strubell, E., Ganesh, A., & McCallum, A. (2020). Energy and Policy Considerations for Modern Deep Learning Research. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(09), 13693-13696. <https://doi.org/10.1609/aaai.v34i09.7123>.

²² Linden, I., Tilman, V. & Laurent, N. (2023). Les techniques d'intelligence artificielle : histoire, développements et défis. *Recherches de Science Religieuse*, 111, 603-624. <https://doi.org/10.3917/rsr.234.0603>

incombe au soignant et au soigné afin qu'ils puissent s'écouter véritablement l'un l'autre et laisser apparaître leurs représentations différentes, voire divergentes. C'est à partir de celles-ci qu'ils pourront penser, décider, donner du sens. La technique doit donc à certains moments être mise de côté pour laisser place à l'attention et à l'écoute : il y a toujours à *faire connaissance* avec celle ou celui que je rencontre, même si – et surtout si – il attend de moi un service.

L'éthique qui se profile ici – que nous pourrions qualifier d'éthique du *oui et non* –, repose donc sur le principe premier de devoir présupposer que l'autre a toujours quelque chose de nouveau à nous apprendre : sa singularité est générique. En médecine, elle peut prendre le nom d'idiosyncrasie, mais étant donné qu'il s'agit d'une singularité vivante, il suffit d'évoquer sa nature continuellement changeante et donc toujours en partie imprévisible, pour comprendre la nécessité de toujours chercher à dépasser les préjugés.

Rappelons enfin ce point sur lequel nous avons déjà insisté : la principale faiblesse des IA, et notamment des IA connexionnistes (Deep Learning), c'est qu'elles sont incapables de donner un sens à ce qu'elles font : tout simplement parce qu'elles sont incapables de faire face à l'inconciliable. Elle restent programmées pour produire une réponse – la plus probable, la plus statistiquement proche d'un optimum – et suivre une seule direction. Or, comme nous le rappelait Heidegger, la pensée humaine (« méditante » et non limitée au calcul) fait constamment face à des interprétations multiples et souvent contradictoires. Elle hésite et doit s'interroger : quel sens choisir, donner, privilégier ? Elle doit dès lors décider et, ce faisant, imposer sa *norme*. De ce point de vue, toute technique demeurera toujours *anormale*. Autrement dit, la technique a besoin de l'humain qui l'utilise pour trouver sa norme, car l'humain verra toujours en elle plusieurs usages possibles : il devra, lui, décider quel finalité convient dans le cadre de la rencontre qui a lieu.
