

REFLEXIONS ET RECOMMANDATIONS SUR L'USAGE DE L'IA EN SANTE

Préconisations du groupe de travail IA

Think Tank Télésanté & Numérique en santé



Texte intégral

Introduction

Le Think Tank Télésanté et Numérique en Santé a constitué un groupe de travail ouvert sur **les usages de l'Intelligence artificielle en santé**. Il ouvre **aujourd'hui une publication séquentielle des travaux** qui ont été menés de septembre à décembre 2025.

Ce groupe n'a pas été composé uniquement d'experts du numérique et de l'intelligence artificielle. Il rassemble autant d'utilisateurs que de concepteurs qui ont réfléchi aux usages des moyens apportés par l'IA en santé dans les pratiques professionnelles comme dans les sollicitations des usagers. On y retrouve des professionnels de santé, médecins et autres praticiens, des ingénieurs des milieux informatiques, des entrepreneurs du numérique, des journalistes, des enseignants et des chercheurs, et des patients. Ainsi ont été mutualisées des réflexions et apports divers de connaissances et compétences.

Nous remercions vivement les apports d'intervenants extérieurs au groupe qui ont apporté leurs connaissances du sujet sous des facettes différentes sur le plan de la stratégie nationale et européenne souveraines, visant au déploiement et à la régulation des usages de l'IA en santé, sur les bénéfices apportés par les usages de l'IA, tout autant que les inconvénients potentiels et les risques, et donc sur la nécessité d'une formation de bon niveau, sur les approches éthiques dès la conception.

Notre groupe de réflexions et de travail s'est fixé comme objectifs de recueillir les connaissances acquises par les professionnels « de terrain » quelques soient les usages qu'ils en ont fait et d'accompagner ainsi les professionnels de santé et les citoyens dans les usages de cette innovation majeure dans le champ du numérique en santé. Ce véritable bouleversement suscite une curiosité active mais aussi inquiète sur les risques encourus dans la relation humaine et humaniste entre un soignant et un soigné, ainsi que sur les risques utilitaires ou financiers qui viendraient altérer les valeurs fondamentales sur lesquelles repose le fonctionnement de notre système de santé.

Notre groupe se fait aussi le porte-voix de « l'expérience terrain » pour proposer une vision orientée sur la médecine basée sur les preuves scientifiques dans ces usages de l'IA, sans négliger la relation soignant-patient. Au cœur des préconisations que nous émettons suite à nos travaux, prédominent les notions de confiance, d'appropriation et d'organisation, à un moment de notre histoire où la théorie du grand remplacement de l'Humain par l'IA fait florès.

Nos réflexions nous ont naturellement engagé à la lecture des nombreux référentiels publics et publiés ces derniers mois (ANS, Ministères, HAS, Académie, et autres Sociétés savantes sans les épuiser toutes, les Associations nationales de patients comme France Assos Santé, etc.). Ce travail s'est aussi appuyé sur une formation interne permettant d'aligner les membres du groupe sur un premier niveau de culture à l'IA, une veille partagée, des lectures croisées et des webinaires thématiques ouverts organisés par le Think Tank et animés par des experts du sujet.

Le groupe de travail a fait part à la DGOS de son intention de contribuer aux travaux et développements de la stratégie nationale et en a informé la DNS.

Conscient de l'accélération technologique et de la diffusion rapide des usages, nous avons choisi d'inscrire nos préconisations autour d'une perspective large sur les points fondamentaux. Une autre étape suivra sur les applications concrètes dans la diversité des pratiques professionnels des métiers de la santé et des nombreuses spécialités. Nous pensons que cette seconde phase devrait s'associer avec les travaux qui sont déjà commencés, ou qui vont l'être, dans les conseils nationaux des différents métiers et des diverses spécialités.

Membres du groupe de travail et les chapitres

Composition du groupe de travail IA en santé
Hervé Pingaud, conduite du groupe Marion Lagneau, responsable animation et rédaction Jacques Lucas, coordinateur communication TT Jerôme Clément Philippe Tronc, Lucile Sivot Sabine Ray-Arles Jean-Yves Mailfert Elénore Scaramozzino Michel Paoli Patrice Ancillon Fabrice Huré Franck Clarot Cédric Villeminot Julie Tamba Thierry Moulin Jérôme Villeminot Thierry Saint Val

Avec le concours de personnalités extérieures au TT, que nous remercions pour leur participation

Dr Yann Maël Le Douarin
Jean Louis FRAYSSE
David Gruson
Clément Goehrs
Dr Alacoque Xavier
Francois Versini

Enoncé des six ensembles de préconisations :

- 1) Voir l'IA comme un dispositif et pas comme un simple outil
- 2) Accompagner la conduite du changement à l'heure de l'IA en santé
- 3) Construire la confiance : comment l'IA a appris, comment raisonne-t-elle, clarté de la place de l'IA entre le soignant et le patient
- 4) Contrôler l'impact environnemental
- 5) Instituer un cadre institutionnel agile
- 6) Appliquer la garantie humaine

Chapitre I

VOIR L'IA COMME UN DISPOSITIF ET PAS COMME UN SIMPLE OUTIL

Enjeux :

L'IA est la révolution technologique la plus rapidement adoptée de toute l'histoire de l'humanité.

Elle résout des problèmes complexes jusque-là réservés à l'humain. Reconnait et interprète le langage naturel, réalise des déductions, des prédictions et des recommandations complexes. Son pseudo langage simule la pensée cognitive de l'humain.

Elle donne aux programmes des facultés proches du cognitif : percevoir, raisonner, agir

>>>> Pour autant, il ne faudra jamais oublier, et il est bon de commencer à le dire, puis à le redire :

- Une IA n'est pas intelligente et n'aura jamais l'intelligence humaine.

- Une IA n'interprète pas la pensée: elle traite de l'information. Elle fonctionne par corrélation et probabilités
- Une IA ne fait pas d'association d'idées.
- Une IA n'a ni inventivité, ni créativité, ni subjectivité . Toutes qualités propres au cerveau humain.

=> **Le statut de l'IA mérite réflexion :**

- Ce n'est pas une intelligence.
- Ce n'est pas non plus un simple outil.
- C'est un service, un « dispositif » ou un « système ».

Ce point conceptuel nous paraît fondamental : on ne doit pas traiter l'IA comme un partenaire de pensée.

Si elle n'est pas une intelligence, quel statut lui conférer ? car elle n'est pas non plus un simple outil.

L'IA N'EST PAS UN OUTIL

Le philosophe Giorgio Agamben distingue trois notions :

- L'outil, qui laisse le choix à l'utilisateur.
- Le dispositif, qui structure les usages et modifie les comportements.
- Le système, qui naît quand plusieurs dispositifs interagissent selon des règles communes.

NON, L'IA N'EST PAS UN OUTIL

- **Un outil est un moyen concret utilisé pour réaliser une tâche précise ou faciliter une activité.** C'est souvent un objet matériel, un instrument technique (physique ou logiciel) qui aide à accomplir une action donnée (par exemple un marteau, un logiciel de rédaction, une application de calcul).
- **Ce n'est pas parce qu'on peut classer l'IA ou définir des niveaux de risque**, des matrices de vigilance opérationnelle ou des panoramas d'usage que cela en fait un outil. (IA symbolique, IA numérique, générative, agentique, machine learning/deep learning pour usages administratif, organisationnel, diagnostic, suivi médical, aide technique, etc...)

EN REALITE L'IA COCHE TOUTES LES CASES DU DISPOSITIF

Un dispositif est structurant, c'est une **organisation ou un assemblage d'outils (techniques, humains, matériels, logiciels)**

L'IA coche toutes les cases du dispositif parce que :

- Elle apprend
- Elle agit en fonction de sa conception, dépendante à 100% des données et des algorithmes
- Elle a un impact transformateur, bien au-delà de son cadre technique
- Elle a des capacités de stratégies et d'agencement complexe des pensées, des discours, des comportements
- Vrai ou faux, peu lui importe, elle désinforme aussi bien qu'elle informe
- Et surtout : l'humain ne maîtrise pas tout

Son pouvoir réel se trouve là: l'IA impose une nouvelle logique d'usage. Elle apporte une nouvelle grammaire de décision. En ajoutant au dispositif des relations, des règles de fonctionnement dans l'idée d'atteindre un objectif commun, elle induit un nouveau système et entraînera des transformations systémiques en santé.

REGARDONS LES FAITS : L'IA SUSCITE DES ATTENTES ALLANT LARGEMENT AU-DELA DE CE QU'ON ATTEND D'UN SIMPLE OUTIL

L'IA semble une bénédiction, susceptible de résoudre de nombreuses contraintes : temps limité, rareté des ressources, charge émotionnelle, épuisement professionnel

- Elle automatise les tâches administratives
- Elle promet une augmentation des connaissances et de l'efficacité clinique, une meilleure efficacité pour recherche de pathologies, synthèse des dossiers, réduction des erreurs .
- En somme, de grandes "Promesses de valeur attendue" : une médecine « plus fiable, moins coûteuse et plus efficace » : plus de temps et de connaissances pour les soignants.

Ce qui se joue va bien au-delà de l'outil. Pourquoi tant de craintes et de préoccupations ?

- Chez les professionnels de santé
 - o La peur de délégation excessive à la machine
 - o La crainte de perte de compétence (ou l'acquisition imparfaite des connaissances et compétences des jeunes)
 - o La crainte que les propositions d'IA influencent trop ceux des humains, parce que ce sont des choix sans conscience, sans émotions, sans jugement, sans connaître ses limites, et sans valeurs morales
 - o Et cette question inavouée, et à poser dès à présent : à quoi servira le temps libéré ?
- Chez les patients, d'autres préoccupations :
 - o Un flou sur la place réelle de l'IA en santé,
 - o Anxiété et inquiétude induites entraînant une multiplication des demandes médicales
 - o Excès d'autodiagnostic particulièrement en psychiatrie, tels le TDAH, déjà confirmés par retour d'expérience et faisant l'objet d'une attention portée par les autorités.

Préconisation [1]: RETABLIR LA HIERARCHIE DES INTELLIGENCES, ADOPTER UNE POSTURE FACE A L'IA

Rétablir la hiérarchie des intelligences

L'IA n'est pas un outil, c'est un véritable dispositif qui agit au sein d'un système transformateur. Elle transforme la cognition, redistribue l'attention, et déplace la responsabilité.

Le risque est clair : à force de déléguer ce que nous savons faire, nous frôlons le risque de déléguer ce que nous *sommes capables de penser*. Or la responsabilité médicale, la décision clinique, la cognition, ne peuvent être déléguées. Décider, est un acte intellectuel et moral, pas un simple calcul. L'humain doit rester le maître

Adopter une posture critique

La prise de décision est un acte humain qui fonde la responsabilité (décider de suivre ou non la recommandation IA). Elle sous-entend d'assumer en toute conscience ses effets

La posture idéale serait de considérer l'IA pour ce qu'elle est : un « assistant virtuel », voire un « apprenti » ou « interne virtualisé » en phase de maturation. Un collaborateur sans conscience. Un système transformateur dont l'humain doit rester le maître et le critique à tout moment.

« L'intelligence, selon Jean Piaget, n'est pas ce que l'on sait, mais ce que l'on fait quand on ne sait pas »

#

Chapitre II

ACCOMPAGNER LA CONDUITE DU CHANGEMENT À L'HEURE DE L'IA EN SANTÉ

Enjeux:

Un principe fondamental: L'IA doit être perçue comme un **appui au raisonnement**, un compagnonnage, et non comme son substitut.

La confrontation à l'anamnèse et aux éléments contextuels est un principe structurant ; elle constitue à la fois un garde-fou et un levier potentiel d'apprentissage. Toute discordance doit être interprétée comme un **signal nécessitant une réévaluation**.

>>>> Les **risques** identifiés dans la littérature et sur le terrain sont réalistes :

- Glissement des compétences cliniques vers la simple validation de proposition algorithmique,
- Perte des savoir-faire techniques et de la maîtrise gestuelle,
- Fragilisation et délitement de la relation thérapeutique modifiée par l'intermédiation numérique.

- Fragilisation de la transmission des savoirs.

>>>> La question du **temps gagné** est centrale et pour le moment sous-exploitée. **A quoi sera utilisé ce temps libéré ?** l'IA permettra-t-elle de redonner toute sa place à la dimension humaine de la médecine et du soin ?

Ces enjeux appellent des Préconisations opérationnelles, organisées autour de trois priorités :

1. Préserver les savoir-faire et les compétences cliniques et techniques.
2. Développer des compétences émergentes spécifiques pour une utilisation éclairée de l'IA.
3. Adapter les méthodes pédagogiques à ces nouveaux défis.

1) PRÉSERVER LES SAVOIR FAIRE ET LES COMPÉTENCES CLINIQUES ET TECHNIQUES

Des **risques en terme de compétence** sont déjà identifiés dans la littérature

- le **“deskilling”** ou perte de compétence des soignants autonomes ou experts : érosion de la relation au patient, perte de compétence clinique et technique
- le **“neverskilling”** ou défaut d'acquisition de compétences des soignants novices : le socle de formation clinique doit être acquis pleinement avant toute délégation à un outil numérique.

Préconisation [2] : Un socle de compétence solide à préserver

- **L'IA en médecine clinique est et doit rester un appui au raisonnement médical, sans s'y substituer.**
- **Le socle de compétences nécessaire à l'exercice du métier doit être acquis par tout médecin (et plus largement par tout professionnel de santé) avant et indépendamment du recours à l'assistance de l'IA.**
- **Il est indispensable de développer et conserver les compétences cliniques analytiques: analyse sémiologique, élaboration d'un diagnostic différentiel, capacité à justifier une décision thérapeutique, lecture non assistée d'examens (imagerie, ECG, biologie).**

Pour accompagner cette Préconisation, il est recommandé de réfléchir à des indicateurs de compétence et de perte de compétence, performants sans être intrusifs, permettant de suivre l'éventuel impact de l'automatisation sur l'apprentissage.

2) DÉVELOPPER DES COMPÉTENCES ÉMERGENTES SPÉCIFIQUES POUR UNE UTILISATION ÉCLAIRÉE DE L'IA

L'usage de l'IA dépasse la compétence technique individuelle: c'est un enjeu comportemental et organisationnel et collectif.

L'interaction homme–IA n'est jamais neutre :

- les interfaces orientent nos comportements par leur design (technologies de design persuasif, c'est-à-dire des interfaces conçues pour influencer nos choix et nos actions).
- nos jugements sont influencés inconsciemment du fait des biais cognitifs induits par l'IA (biais de confirmation, biais d'autorité, effet d'automatisation, etc.).

Le risque d'une utilisation individuelle sans partage d'expérience ni analyse collective des erreurs est un point clé souvent négligé

La seule solution: L'usage de l'IA en santé doit devenir un usage collectif, documenté et discuté, et non une pratique individuelle invisible

Des compétences opérationnelles doivent émerger, centrées sur :

- La compréhension des concepts, des limites, des biais et du caractère probabiliste des modèles d'IA, notamment générative.
- Le développement d'une vigilance critique active, plutôt que d'une confiance aveugle dans les outils.

Des compétences clés devront être développées :

- Maitrise l'art du « prompt » (requête) pertinent : la formulation conditionne directement la qualité de la réponse. Seule une anamnèse rigoureuse et structurée alimentera correctement le prompt.
- Exercer une vigilance continue face aux **biais, erreurs et hallucinations**, qui peuvent se traduire par des assertions fausses, mais formulées avec assurance
- **Superviser** la compréhension des données et des méthodes de validation et de la qualité des données plutôt que la logique algorithmique,
- Adopter en toute circonstance une **posture critique et argumentée** vis-à-vis des résultats fournis par l'IA

Préconisation [3] : Des compétences spécifiques à acquérir

La formation à l'IA doit s'axer vers des compétences spécifiques (art du prompt, par exemple), mais aussi sur la vigilance de l'esprit critique et la qualité d'interprétation des résultats,

Une approche équilibrée entre préservation des savoir-faire humains et acquisition de compétences IA est obligatoire

Cette approche devra être intégrée aux enseignements fondamentaux de tous les métiers de la santé (médecins, infirmier·e-s, pharmaciens, sages-femmes, kinésithérapeutes, etc.), en formation initiale, comme en formation continue.

3) ADAPTER LES MÉTHODES PÉDAGOGIQUES À L'ÈRE DE L'IA.

Pour répondre à l'innovation apportée par l'IA, il faut **une pédagogie à la hauteur des enjeux**.

L'objectif est de construire une culture commune de l'IA en santé, fondée sur la compétence, la responsabilité et la vigilance collective.

- **Modules interactifs et ludiques:** cas pratiques, modules courts, ateliers de simulation, jeux sérieux, vidéos.
 - o **Objectif:** Apprendre à décoder le design persuasif des interfaces (captologie), à repérer les biais et à construire une posture critique, et interdisciplinaire
- **Apprentissage par l'erreur:** utiliser les mises en défaut de l'IA (erreurs, hallucinations) comme support pédagogique pour développer la culture de vigilance critique
- **Interfaces lentes et frictions cognitives:**
 - Introduire volontairement des étapes de réflexion supplémentaires pour restaurer la réflexion clinique.
 - *Exemple concret : avant de valider une proposition de diagnostic ou de traitement générée par l'IA, imposer une justification clinique argumentée par le soignant, ou une vérification obligatoire de certains éléments clés (signes d'alerte, contre-indications, alternatives).*
- **Encadrement et usage collectif de l'IA:**
 - o **Double lecture humain + IA**, notamment en imagerie, dermatologie ou anatomo-pathologie
 - o Mise en place de dispositifs de traçabilité, d'indicateurs et d'audits (taux de relecture, erreurs IA recensées, écarts IA/humain).
 - o Renforcement de la collégialité : discussions entre pairs, retours d'expérience, comités IA pluridisciplinaires (intégrant médecins, soignants paramédicaux, data scientists, représentants des patients...).
 - o Constitution de registres ou bases de cas où l'IA s'est trompée ou a posé problème, pour permettre un apprentissage collectif.
- **Référents IA dans les sociétés savantes**, par spécialité, et par professions, et dans les ordres professionnels
 - o afin de favoriser les relations entre pairs et les temps d'analyse collégiale
 - o Les ateliers, les conférences et les discussions facilitent l'adhésion des parties prenantes
- **Dispositifs nationaux d'acculturation pour l'ensemble des professionnels de santé**
 - o Développer des certifications spécifiques (par exemple, PIX Pro Santé) et renforcer les modules IA proposés par l'ANS (actuellement trop peu nombreux).
 - o Intégrer de manière structurée l'usage de l'IA dans les enseignements fondamentaux des métiers de la santé.

Le décret actuel prévoyant 28 heures de formation aux outils numériques durant les études médicales est jugé très insuffisant, tant en volume qu'en objectifs pédagogiques

Préconisation [4] : Une pédagogie innovante !

L'essor de l'IA en santé appelle une évolution des modes de transmission des savoirs en santé.

L'IA en santé progresse vite, parfois plus vite que notre capacité collective à la comprendre. Pour éviter les incompréhensions et les usages inadaptés, une pédagogie innovante autour de l'IA en santé n'est plus une option : elle s'impose comme une condition essentielle et une nécessité pour accompagner ces transformations et doit faire l'objet d'une réflexion rapide et structurée.

Engager cette réflexion sans attendre est un enjeu majeur pour l'avenir du système de santé.

Chapitre III

CONSTRUIRE LA CONFIANCE : DU SOIGNANT UTILISATEUR AU PATIENT QUI S'INFORME

- **Avoir confiance dans l'IA utilisée : comment apprend-t-elle ? comment raisonne-t-elle ?**
- **Avoir confiance dans les réponses obtenues : avoir réponse à tout est-il synonyme de fiabilité ?**

Enjeux :

La collaboration en santé repose sur la confiance dans le « référentiel » du « collaborateur » : connaissance de la qualité de son apprentissage, sa méthode de raisonnement, surtout si cela implique une proposition de diagnostics ou de thérapeutiques. Avec une IA, même exigence. Pas de confiance aveugle. L'IA doit être vue comme un **outil d'aide à la décision intégré dans une chaîne de responsabilité**. Le soignant reste responsable de la décision et doit être clair avec le patient: qui le soigne ? et sur quels critères ?

1) **BRIQUE DE CONFIANCE NUMERO 1 : Connaître “la boîte noire” de l'IA qu'on utilise**

L'opacité du fonctionnement interne d'une IA s'apparente à un effet “boîte noire” : **ce n'est pas parce qu'une réponse est convaincante qu'elle est fiable.**

il est indispensable d'avoir des états d'âme et ne pas accepter ses capacités sans recul.

Quelques exemples de questions essentielles avant de faire « confiance » à une IA

- Quel type d'IA utilise-t-on : générative, symbolique, autre ?

- Sur quoi a-t-elle appris, et à partir de quelles sources ? Permettre au médecin de comprendre les principaux éléments qui orientent la suggestion.
 - Comment a-t-elle été entraînée ? : données limitées ou homogènes ? couvre-t-elle toutes les catégories de patients, formes cliniques et contextes de soin ?
 - Peut-on suivre son raisonnement ? La fiabilité ne dépend pas de l'algorithme, mais de la composition des données qui l'ont façonné.
 - Peut-elle être fidèle à une "école de pensée" particulière ?
 - Quelles protections contre la désinformation et les fausses informations ? Mesures intégrées au modèle linguistique ? Procédure de signalement des défauts ? Mécanismes de correction ?
- ⇒ **Sans réponses à ces questions, le risque est présent:** résultats biaisés, analyses trompeuses, performances inconstantes, conclusions inexactes.

Préconisation [5] : Connaître la boîte noire

=> La confiance n'exclut pas le doute

Utiliser l'IA à bon escient, c'est garder une distance suffisante pour rester vigilant.

- **C'est avoir une connaissance de ses mécanismes, et la maîtrise de son fonctionnement**
- **Sans jamais abdiquer son esprit critique,**
- **C'est accepter de parler de risque,**
- **C'est accepter de douter de ses réponses, le doute est la matière première de la mise en confiance. Il doit exister, mais aussi persister ! : la confiance doit être maintenue, contrôlée et maîtrisable dans le temps.**

2) BRIQUE DE CONFIANCE NUMERO 2 : Explicabilité et transparence pour un usage en santé

On « adopte » une IA en santé sous conditions strictes d'explicabilité et de transparence.

- Qualité et diversité des sources et des données : pilier d'adoption numéro 1
 - Comparabilité possible et facile aux autres IA avec des critères partagés : explicabilité, fiabilité, charge cognitive, niveau de délégation...
 - Traçabilité claire des versions et des mises à jour,
 - examen humain préalable,
 - note d'impact clinique,
 - numérotation des versions obligatoires avec date, performance, périmètre.
- Aujourd'hui, des IA peuvent se mettre à jour automatiquement sans validation médicale.
- Préséance humaine - non négociable - à toutes les étapes de conception, validation, usage
 - Contrôles réguliers : détection d'hallucinations, de biais, de dérives

- Tests de reproductibilité (même prompt => constance de la réponse avec une réponse stable à défaut d'être identique)

=> **Une voie prometteuse pour renforcer la confiance :**

Combiner IA générative et IA symbolique (basée sur des règles explicites),

L'idée : développer des contrôleurs (IA symbolique) qui viendront inspecter en temps réel les résultats produits par une IA numérique générative, pour en améliorer la lisibilité et la transparence

Préconisation [6] : Éclairer la boîte noire

- Envisager des contrôleurs d'IA symbolique pour éclairer la boîte noire de l'IA générative .
- Créer un indice d'acceptabilité professionnelle (IAP) porté par des sociétés savantes
 - il serait légitime, structurant et très utile pour orienter les choix d'outils dans les équipes
 - simple et mesurable, utilisable par tous ,
 - facilement mis à jour,
 - permettant de comparer les outils sur lisibilité, explicabilité, fiabilité perçue, charge cognitive induite, niveau de délégation accepté, Ecoscore

3) BRIQUE DE CONFIANCE NUMERO 3 : Le patient doit avoir confiance dans l'IA utilisée par le professionnel de santé

Point fondamental : l'IA n'examine pas, ne diagnostique pas et ne décide pas à la place d'un médecin

La question « *Qui me soigne ?* » doit trouver une réponse claire : l'IA n'est qu'un appui, la décision reste médicale.

Le patient doit pouvoir refuser l'utilisation d'une IA sans que cela altère la qualité du soin (attention: sauf cas particulier d'outil indispensable à la sécurité comme un dispositif médical implantable par exemple).

Préconisation [7] : Informer clairement le patient

- Le niveau d'utilisation de l'IA dans la décision soignante doit être clair et compris par le patient
- Intégrer dans le code de déontologie l'obligation d'informer sur l'usage d'une IA dans la décision de soin proposée.
- Inscrire l'information sur l'IA dans les recueils de consentement et les supports d'information
- Cette transparence contribue à prévenir la défiance et à maintenir un cadre éthique compatible avec l'autonomie, la bienfaisance, la non-malfaisance, et la justice

4) BRIQUE DE CONFIANCE NUMERO 4: l'IA que le patient questionne

Les patients utilisent de plus en plus des outils d'IA.

Leurs connaissances sur ces outils restent limitées, tout comme celles des professionnels: dans une étude, le niveau de connaissance était estimé à 2,4/10 chez les patients, 3,8/10 chez les professionnels.

Connait-on les attentes d'un patient questionnant une IA sur sa santé ?

- Un appui, un complément ou un substitut aux professionnels de santé ?
- Une présence continue?: par exemple en santé mentale, certains systèmes (thérapie cognitivo-comportementale dispensée par chatbots, thérapie d'exposition en réalité virtuelle) offrent une assistance continue sans les contraintes de la fatigue humaine..
- Une perception d'une absence de jugement ? Des systèmes comme ChatGPT montrent une capacité à reconnaître et exprimer des émotions, avec des performances parfois comparables ou supérieures à la conscience émotionnelle moyenne de la population

Quels sont les risques pour les patients questionnant une IA ?

- **L'absence de garantie** : L'IA parle de bien-être et pas de santé : elle n'a pas conscience de ses limites et ne saura jamais donner du secours à une personne dans la détresse
- **Ne consulter que l'IA** sans aller voir un médecin: L'IA peut-elle être suffisante pour un patient sans professionnel de santé ?
- Obtenir une **liste brute d'hypothèses diagnostiques**
 - qui inquiètent
 - entraînent une surconsommation de bilans et de soins
 - peuvent générer une forte anxiété, voire des drames, en particulier en santé mentale (suicide)
- **Attendre trop** de l'IA, comme si c'était une intelligence supra-humaine
- **Se couper progressivement du réel**, capté par l'information apportée par la seule technologie.
- **Mésinformation et désinformation massives en santé**. Les réseaux sociaux font partie des sources des IA grand public. Or...

TikTok présente la plus haute prévalence de désinformation, avec environ 20 % des contenus contenant des fausses informations ; Facebook et X suivent avec respectivement des taux de 13 % et 11 %. 43% des posts considérés comme de la mésinformation sur les réseaux sociaux concernent la santé. En 2022, 51% des posts concernant les vaccins sur les réseaux sociaux contenaient de la mésinformation.

Préconisation pour les patients [8] : douter, vérifier, contextualiser

- En tant que patient, à l'égal des professionnels de santé = douter.
- L'IA n'est pas un service de santé numérique simple et sécurisé, centré sur l'utilisateur.
- Connaître l'ampleur de la mésinformation et de la désinformation sur les réseaux.
- Ne pas surestimer les capacités de l'IA : réaliser que les applications actuelles dans le domaine de la santé s'apparentent, en l'état, à des expériences en conditions réelles menées sans consentement explicite.
- Ne pas ignorer les intérêts économiques sous jacents

La confiance dans l'IA au service du soin impose doute raisonné, vigilance, transparence, et présence humaine.

Chapitre IV

CONTRÔLER L'IMPACT ENVIRONNEMENTAL

Enjeu :

L'empreinte environnementale de l'IA, soulignée par le rapport du **Shift Project**, ne doit pas être occultée. Les coûts énergétiques et environnementaux sont considérables.

IA et énergie : quelques points clés

Aucun grand acteur d'IA ne fournit de données détaillées, transparentes et fiables sur sa consommation énergétique, et donc son empreinte GES (gaz à effets de serre).

On a des ordres de grandeur, pas des bilans complets.

Quelques repères simples :

- une requête sur ChatGPT = consomme environ 2 wattheures (Wh) d'électricité
- une recherche sur Google = 0,3 Wh
- ⇒ soit environ 6 fois plus pour une requête IA, du fait des milliards d'opérations nécessaires.
- ⇒ Plus la requête est complexe, plus elle consomme : La création d'une image en haute définition par une IA consomme autant d'énergie que la recharge complète d'un téléphone portable.

Au-delà de l'impact énergétique : l'impact environnemental

Il commence bien avant la première requête. Le numérique est une industrie lourde, et contrairement à l'impression ressentie, elle n'est pas dématérialisée, au contraire, elle est très matérielle.

- La fabrication des serveurs utilise des métaux rares.
- Les data centers consomment beaucoup d'eau, pour climatiser et refroidir. Dans certaines régions, cela induit une concurrence avec l'eau potable locale et un stress hydrique.
- Une grande partie de l'énergie utilisée pour produire l'électricité nécessaire provient encore de sources non renouvelables comme le charbon, le pétrole et le gaz.

Impact énergétique d'une requête IA : quelques chiffres et comparaisons

Les calculs en IA générative utilisent un réseau de neurones profonds pour copier le fonctionnement du cerveau humain. Les calculs sont assurés par des processeurs GPU (Graphics Processing Unit) spécifiques dédiés dont la puissance électrique consommée est considérable.

La consommation totale d'une requête IA se décompose en **deux séquences** :

- L'énergie utilisée pour la requête : elle dépend du modèle de puce utilisé, du nombre de requêtes, du moment de la requête (pics), de la complexité de la question.
- L'overhead du datacenter : c'est la partie la plus consommatrice d'énergie (refroidissement, alimentation, réseaux, pertes électriques)

Répartition de la consommation énergétique d'une requête :

- **La requête** : un grand modèle de langage (LLM) consomme: → 0,2 à 0,6 Wh de calcul brut (valeur courante mesurée : 0,3 Wh)
- **Le volet fonctionnel du centre** => entraîne l'explosion de la consommation du fait de: quantité d'eau utilisée, dégagement de chaleur, refroidissement. Refroidir les serveurs est le vrai gouffre énergétique.

L'indicateur clé actuel est le PUE (Power Usage Effectiveness). Il mesure la part d'énergie consommée par le datacenter par rapport à celle utile au calcul. Il traduit l'efficacité énergétique du site.

Un PUE de 2, c'est comme une voiture qui consommerait autant d'essence pour avancer que pour se refroidir (1 KWh pour la requête + 1 KWh pour refroidir)

La **même requête IA** (par exemple une requête de génération de texte ou d'image) ne consomme pas la même énergie selon le data center : **un datacenter performant peut traiter une requête avec presque deux fois moins d'énergie qu'un datacenter mal optimisé**

En pratique, le PUE reste souvent supérieur à 1,5, malgré les progrès des grands acteurs comme Google ou Microsoft qui visent le $PUE \leq 1,1$ dans leurs infrastructures les plus modernes.

Bilan :

- **Une même question d'IA, selon le data center, c'est comme si on l'exécutait dans un bâtiment bien isolé... ou dans une passoire thermique**

- **Le cerveau humain est capable de faire la même chose qu'une IA, et sans s'échauffer. L'IA ne remplacera pas le cerveau humain (contrairement à ce que certains croient), elle économisera son fonctionnement (ce qui n'est pas sans risque, voir chapitre II, la conduite du changement face à l'IA). En revanche, l'IA n'est pas partie pour économiser l'énergie de la planète :**
 - ⇒ **Si on compare une IA au cerveau humain, pour un même calcul, la puissance consommée par une IA est environ 25 000 fois plus grande !**

Comment appliquer une dose de sobriété numérique en santé

En tant qu'utilisateur : c'est avant tout avoir un esprit critique vis-à-vis de l'usage de l'IA

- Ne pas se rassurer en se disant qu'une requête individuelle est faible par acte. Ce qui compte, c'est le nombre immense d'actes, et donc le volume cumulé d'une utilisation massive.
- Toujours se poser la question clé : que permet-elle d'éviter ou d'améliorer, et pour quel coût énergétique ?
- Utiliser l'IA
 - quand il y a un bénéfice clinique identifié,
 - dans les situations où elle améliore réellement le soin
- Favoriser des modèles adaptés (pas surdimensionnés), et des usages mutualisés

A l'échelon de la planète :

- **Diminuer le gaspillage d'énergie :** les IA actuelles ont une pluralité d'usage, et utilisent trop de ressources pour des requêtes simple. Cet axe doit être optimisé, avec une autre approche, la frugalité.
- Des infrastructures **efficaces énergétiquement**.
 - Implantation **géographique** des data centers réfléchies en fonction des **conditions climatiques locales** (les zones tempérées sont à privilégier) et en utilisant de l'électricité d'origine renouvelable (énergie solaire, éolienne, hydroélectrique, géothermique)

La France dispose d'atouts réels pour l'implantation de datacenters :

- position géographique stratégique
- électricité majoritairement décarbonée, ce qui constitue un avantage environnemental important.
 - ⇒ C'est un avantage environnemental majeur pour l'implantation de datacenters liés à la santé.

Enfin, on recommande la même sobriété pour l'IA que pour les prescriptions médicales !

- « *Utiliser une IA trop complexe pour une tâche simple juste parce qu'on l'a sous la main, c'est comme prescrire un scanner pour explorer une piqure de moustique* ».

- « L'IA c'est comme les antibiotiques : le problème n'est pas la prescription, mais la bonne dose, dans la bonne indication et sans usage inapproprié ».

Préconisation [9] : inscrire la sobriété numérique dans le pilotage de l'IA

- **Intégrer la sobriété numérique et une évaluation environnementale dans les cahiers des charges, conformément à la loi « climat et résilience »**
- **L'enjeu est d'utiliser l'IA en santé là où son bénéfice clinique dépasse son coût énergétique**
- **Privilégier des modèles frugaux, ayant une dépense énergétique adapté à la tâche demandée, et mutualiser les infrastructures.**
 - **C'est déjà l'esprit du RGPD qui impose de limiter les données traitées à ce qui est strictement nécessaire au résultat attendu**
- **Exiger que les Conditions Générales d'Utilisation d'une application d'IA informent clairement l'utilisateur sur l'empreinte environnementale de l'outil.**
- **Rendre obligatoire une évaluation du rapport bénéfice/impact carbone dans les projets et appels d'offres IA, en application de la loi « climat et résilience » de 2021. Par exemple, un "éco-score" automatique associé à chaque requête (type ticket de caisse carbone)**

Pour approfondir

Datacenter efficace ou énergivore : les différences de PUE

Quelques repères de PUE actuels

PUE (Power Usage Effectiveness)

- Datacenter très efficace (Nord Europe, sites neufs, refroidissement naturel) : PUE $\approx 1.1 - 1.2$
- Datacenter standard (Europe de l'Ouest / USA récents) : PUE $\approx 1.3 - 1.5$
- Datacenter ancien ou en climat chaud (USA sud / Asie / sites à refroidissement mécanique) : PUE $\approx 1.6 - 2.0$

- Les plus économes se trouvent dans les pays nordiques (climat frais, refroidissement naturel, récupération de chaleur), Irlande mise à part.
- Les plus coûteux sont les datacenters anciens ou situés dans des climats chauds (Dubai, Texas, Inde), où le coût du refroidissement explose.

Référence (entre autre) : <https://drane-versailles.region-academique-idf.fr/spip.php?article1167>

Chapitre V

INSTITUER UN CADRE INSTITUTIONNEL AGILE

Enjeux

L'IA va vite, très vite, parfois trop vite...

- Le développement numérique est tellement rapide qu'il impose souvent des normes a posteriori, basées sur la force des usages.
- C'est donc une utopie de penser qu'il y aura une « loi universelle qui s'imposerait à tous ».
- Il y a des bases posées dans L'AI Act de 2024, mais chaque solution proposée soulève des questions pratiques : « Qui supervise ? Qui paie ? ».

Le cadre institutionnel

1) Adapter le cadre réglementaire

- Et vite ... dans des fenêtres de temps réalistes vis à vis de la dynamique de l'offre et de la demande (*en France à ce jour, environ 350 sociétés de santé numérique intégrant au moins un module ou usage de l'IA et santé et autour de 100 start-up IA dont le principal cas d'usage est la santé*)
- Le règlement de **l'IA Act** -règlement européen fixant le cadre de base pour l'IA (risques, obligations, interdictions)- semble déjà imposer trop de charges réglementaires.
- Un autre projet de règlement européen, **Digital Omnibus** propose de modifier l'IA Act pour rendre l'écosystème plus lisible, tout en restant compatible avec le « modèle européen » de protection des données pour que les entreprises n'aient pas à gérer des obligations qui se chevauchent.

2) Ne pas ajouter de la (léendaire) rigidité administrative à la française

- Règles de création de start up, d'implantation de data center : la célérité de l'IA ne va pas avec la lenteur réglementaire.
- Bien évidemment, ce travail devant se faire, à la fois par rapport aux règles existantes du système de santé en y apportant les adaptations nécessaires pour ces nouvelles technologies.
- Tout en gardant une vraie prudence face à la puissance de certaines nations, de gros éditeurs et d'intérêts commerciaux

3) De la concertation entre les agences d'état, stop au travail en silo

- Le travail normatif devra forcément être collaboratif : de nombreux acteurs étant impliqués, (professionnels de santé, utilisateurs des dispositifs médicaux défendant les conditions de leur activité, et industriels qui les développent),

4) L'implication des sociétés savantes, et du corps des enseignants et universitaires, des différents conseils nationaux des spécialités ou des disciplines.

- La production de recommandations par la profession elle-même est un vecteur de diffusion des bonnes pratiques. Les questions de responsabilité professionnelle et de responsabilité civile nécessitent une conception et une mise en œuvre minutieuses afin d'éviter les conséquences imprévues et doivent être abordées

Le cadre éthique

- Les travaux en éthique de l'IA se sont multipliés.
- L'implication systématique des patients dans la gouvernance (co-design) est un standard éthique en croissance (international, FDA, NICE...).
- Le cadre légal européen, incarné par l'AI Act de 2024, est jugé nécessaire mais insuffisant du fait que sa mise en œuvre s'étale jusqu'en 2028. Sa déclinaison sur le terrain opérationnel focalise surtout les éditeurs et assez peu sur les usagers.

Quelles solutions pourraient faciliter la gouvernance collaborative ?

- 1) **Un observatoire national ou une cellule de veille** pour exercer un contrôle, suivre l'évolution de l'offre et permettre des retours d'expérience structurés des usages par des pilotes mis sous surveillance et dans des délais raisonnables. C'est la transposition du principe de **bac à sable** de systèmes d'IA de l'IA Act qu'il faut rapidement incarner au sein d'un tel observatoire.
- 2) Des **“Lieux d'Expérimentation Éthique”** (Living Labs IA Santé) : espaces où soignants, patients, développeurs et régulateurs testent les IA en conditions réelles avant tout déploiement large. Leur valeur ajoutée : évite les déploiements massifs sans validation humaine.
- 3) **La mise en place d'un « label de transparence » et une « labellisation minimale »** des outils d'IA, assortie de sanctions face aux risques, et d'une périodicité de renouvellement adaptée (processus continu et potentiellement heuristique)
- 4) **Le référencement des solutions IA** avec l'évaluation + d'autres critères à définir, sur une plateforme dédiée aux professionnels de santé.
- 5) **Un encadrement précis des IA “conversationnelles”** destinées aux patients mais aussi aux aidants. Les approches relatives au consentement des patients et au partage d'informations nécessitent une reconnaissance explicite du caractère expérimental de l'IA.
- 6) **Un renforcement des CGU** (modèle, version, mise à jour, responsabilités) : L'utilisateur devant pouvoir choisir ses outils IA, le LLM mis en œuvre et son niveau de version, les dates de mises à jour des algorithmes, open source ou pas, des données pour le suivi du niveau des connaissances embarquées, une description claire des responsabilités des productions (sous-traitant, tiers technique, fournisseur des données, responsabilité des traitements, des données, ... en plus de ce que le RGPD oblige déjà à décrire.

- **Préconisation 10 Privilégier une approche de « droit souple » (« soft power ») : « Influencer plutôt que contraindre »**
- **Décloisonner la gouvernance des organismes d'État, mettre en place une collaboration d'acteurs engagés autour d'une gouvernance centralisée et restreinte. Les acteurs des concertations doivent être non seulement les éditeurs, mais également les usagers**

- Des capacités de mise en application rapide, tout en gardant des garde-fous coercitifs si nécessaire
- Un cadre légal et éthique solide, faisant intervenir les éditeurs mais aussi les usagers, est indispensable : il doit pouvoir donner aux utilisateurs une capacité éclairée de choix de la solution utilisée

Chapitre VI

APPLIQUER LA GARANTIE HUMAINE

Enjeux

Les applications de l'IA en santé demandent non pas à être entravées mais régulées par **deux principes éthiques** organisateurs :

- Un principe de **fraternité** envers les personnes vulnérables : on évitera au maximum de remplacer les métiers du soin (infirmiers, psychologues cliniciens, etc.) par des robots humanoïdes dotés d'IA. La fraternité témoigne de notre attachement à l'authenticité des échanges et à la dignité des personnes,
- Un principe de « **garantie humaine** » : le patient doit pouvoir toujours compter sur la présence d'un médecin.

Parmi les facteurs clés de succès qui se dégagent dans tous les contextes, **on peut citer le maintien d'un contrôle humain associé à l'autorité de décision finale**: Ce principe de garantie humaine, devenu une norme AFNOR (et présent dans la norme internationale ISO/IEC 42001), est essentiel pour superviser l'IA à toutes les étapes de son cycle de vie et évaluer les « effets indésirables » qu'elle peut générer.

La France et l'Europe ont posé le principe de « garantie humaine » à l'interprétation des résultats diagnostiques et thérapeutiques, impliquant que la machine ne peut agir seule et consacrant la responsabilité du professionnel de santé.

Préconisation 1 1

Il faut exiger une “Traçabilité Humaine” dans toute décision assistée par IA.

La fraternité et la dignité impliquent de ne pas substituer les métiers du soin par des agents artificiels dans les situations relationnelles critiques.

Le principe de « garantie humaine » impose que la décision médicale dans le parcours de soin reste toujours la décision humaine.

Métaphoriquement, nous tenons à rappeler que l'humain doit garder le premier et le dernier kilomètre de la réflexion.