

# Intelligence artificielle et santé mentale

Définir les frontières entre  
innovation, science et éthique



Octobre 2024

**MentalTech** × **stane**

# Sommaire

<b>À propos</b>	p 3
<b>Préambule</b>	p 4
<b>Méthodologie</b>	p 8
<b>I. Les quatre cadres d'intervention de l'IA en santé mentale</b>	p 10
<b>II. Les usages concrets de l'IA en santé mentale</b>	p 12
<b>1. Diagnostic et évaluation</b>	p 14
Détection et diagnostic de troubles mentaux	p 14
Aide à la décision clinique	p 15
Évaluation des risques et pronostic	p 17
<b>2. Traitement et suivi</b>	p 18
Suivi et monitoring	p 18
Thérapie et réhabilitation	p 18
<b>3. Amélioration des pratiques et de l'accès aux soins</b>	p 21
Amélioration de l'accès aux soins	p 21
Gestion du stress et du bien-être	p 22
Formation et supervision des thérapeutes	p 22
Prévention des crises	p 23
Recherche et développement en santé mentale	p 24
<b>III. La réglementation en vigueur concernant l'IA peut-elle couvrir l'intégralité des risques identifiés ?</b>	p 25
<b>IV. Quelles préconisations pour atténuer les risques inhérents à l'utilisation de ces Intelligences Artificielles ?</b>	p 28
<b>Conclusion</b>	p 32
<b>Les contributeurs</b>	p 35

# MentalTech

Le collectif MentalTech est le premier collectif français dédié à l'émergence de solutions numériques en santé mentale. Il est créé en mars 2022 par 7 membres fondateurs, rejoints depuis par une trentaine d'acteurs : institutionnels, startups et professionnels de santé. Il a vocation de rassembler organismes privés et publics autour d'une ambition forte : répondre à l'urgence de déployer des outils numériques éthiques dans la prévention et la prise en charge de la santé psychique.

La présidence de ce collectif est incarnée par le Dr. David Labrosse, médecin de santé publique de formation et fondateur de Tricky, une solution de prévention primaire immersive et ludique destinée aux entreprises et aux patients.

[www.mentaltech.fr](http://www.mentaltech.fr)

## stane

**stane** est une société à mission engagée dans l'amélioration de la santé de tous et la pratique des soignants.

Que vous soyez professionnel de santé, porteur d'un projet innovant en santé cherchant à accélérer votre accès au marché, ou évaluateur de produits de santé, **stane** offre une gamme complète de services, allant de l'accompagnement dans la recherche clinique à la co-conception et l'évaluation de produits innovants, en passant par le déploiement de programmes de recherche en médecine de ville. Avec **stane**, bénéficiez d'une approche facilitée, rapide et stimulante pour améliorer vos pratiques, développer vos innovations ou mener vos études cliniques, le tout de manière efficace grâce à notre large réseau d'experts et de professionnels de santé investigateurs.

Parce que la recherche et l'innovation doivent se déployer en ville, rejoignez nous, soyez acteur du changement : Ensemble, innovons pour la santé !

[www.stane-groupe.fr](http://www.stane-groupe.fr)

# Préambule

La question de la santé mentale suscite une préoccupation croissante, comme en témoigne une récente étude de Santé publique France<sup>(1)</sup>. Cette dernière met en évidence une détérioration persistante de la santé mentale des Français, particulièrement marquée chez les jeunes adultes âgés de 18 à 24 ans.

Les statistiques alarmantes révèlent une recrudescence significative des troubles de l'humeur, des idéations suicidaires et des épisodes dépressifs. En outre, on constate une inquiétude grandissante quant à l'augmentation des consultations aux services d'urgence liées à ces problématiques depuis 2021.

Plus que jamais en ce XXI<sup>e</sup> siècle, dans un environnement climatique, sanitaire et économique en perpétuelle mutation, la santé psychique s'avère indispensable à l'épanouissement de l'être humain et de la société. Elle lui permet de s'adapter et de s'épanouir, tant sur le plan individuel que collectif.

Les outils numériques représentent une opportunité sans précédent pour notre santé mentale, et notre santé de manière générale. Ils démocratisent l'accès à la prévention, favorisent l'autonomie des personnes concernées, optimisent la relation thérapeutique et rendent, in fine, notre système de santé plus efficient.

L'accès aux soins est un aspect crucial des avantages apportés par les outils numériques dans le domaine de la santé. Ce bénéfice est particulièrement visible avec l'émergence des plateformes de téléconsultation, dont Qare est un exemple notable en France : faciliter les consultations pour les personnes vivant dans des zones rurales ou des déserts médicaux, offrir des rendez-vous rapides pour des problèmes de santé non urgents, permettre aux patients à mobilité réduite de consulter sans se déplacer, proposer des créneaux de consultations en dehors des heures traditionnelles, etc.

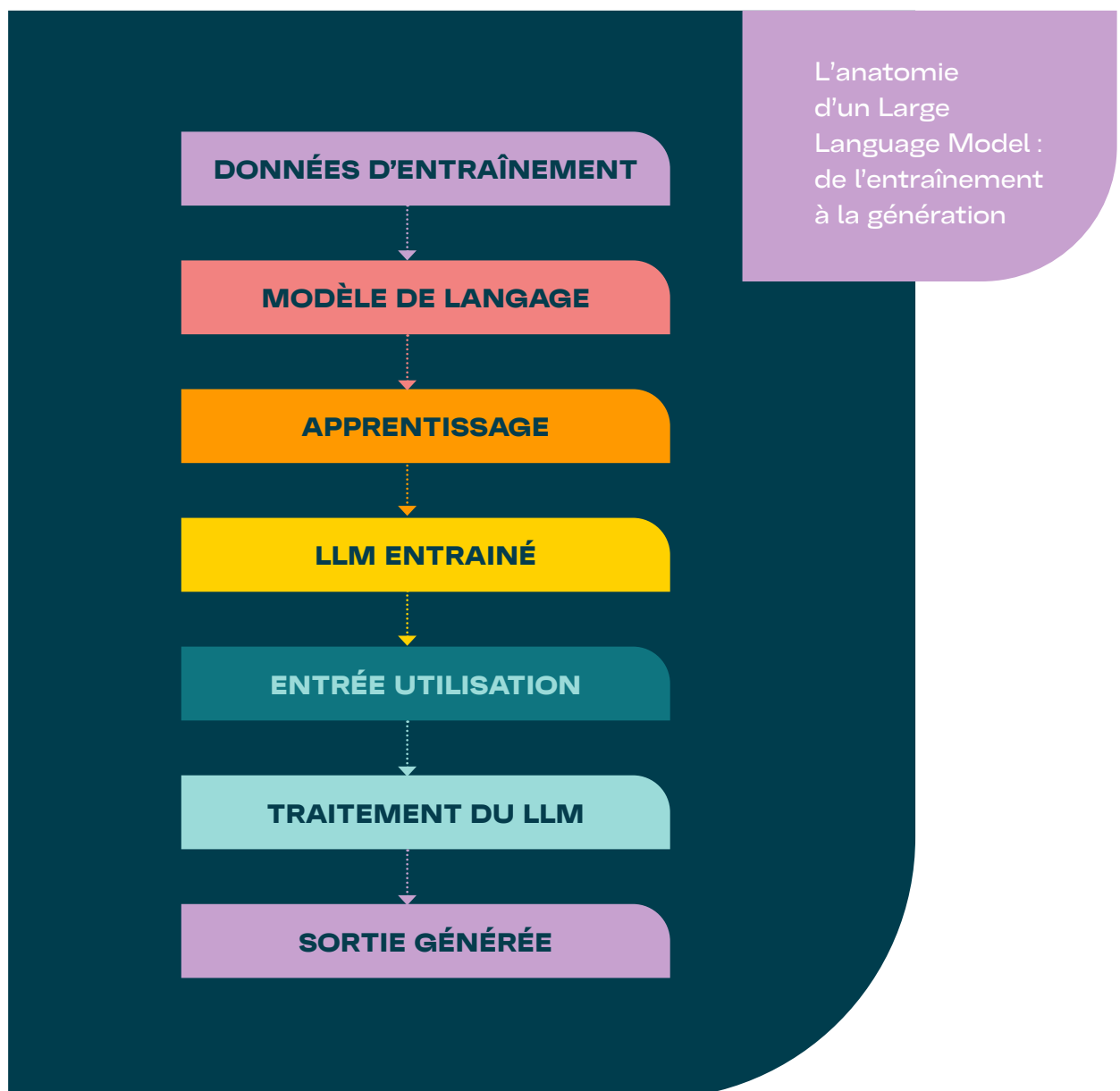
Les utilisateurs de ces nouvelles solutions, adaptées à leur culture et conçues de manière éthique, doivent pouvoir s'appuyer sur ces dernières en toute confiance. Leurs concepteurs se doivent de veiller à l'utilité publique et au respect de la souveraineté de la France en matière de données comportementales et de santé psychique.

---

(1) <https://www.santepubliquefrance.fr/presse/2023/sante-mentale-des-jeunes-des-conseils-pour-prendre-soin-de-sa-sante-mentale>

Au sein du collectif MentalTech, nous sommes convaincus que le numérique n'est plus une option, mais une nécessité pour répondre aux usages de la population, et réduire l'écart entre leurs besoins et l'offre de santé actuelle.

Au cours des dernières années, l'intelligence artificielle (IA) a connu des avancées technologiques considérables, accroissant significativement ses capacités et ses applications dans divers domaines. L'un des aspects les plus notables de cette progression est l'émergence et la popularisation des modèles de langage de grande taille (LLM, pour Large Language Models), tels que GPT (Generative Pre-trained Transformer).



La démocratisation de l'IA, notamment par l'intégration de ces technologies avancées dans des produits et services accessibles au grand public, a facilité une meilleure compréhension et une plus grande acceptation de l'IA par la société. Cette acceptation croissante est également soutenue par les améliorations continues en matière de facilité d'utilisation, de fiabilité et d'éthique de l'IA, renforçant ainsi la confiance des utilisateurs et élargissant les perspectives d'application de l'IA dans un nombre croissant de secteurs, notamment celui de la santé.

En intégrant plus spécifiquement l'IA dans les stratégies de soins de santé mentale, nous pouvons sensiblement améliorer l'efficacité et la qualité des soins proposés. Cela permet de rendre le soutien plus personnalisé, accessible et orienté vers la prévention.

## **Examinons comment l'IA peut contribuer à l'amélioration de la santé mentale.**

Tout d'abord, ces technologies offrent une analyse objective des données pour **accroître la précision des diagnostics**. Cette capacité dépend de la qualité des données sur lesquelles les systèmes d'IA ont été entraînés. Les LLMs ou d'autres modèles d'IA comme les réseaux de neurones convolutifs (CNN) utilisés pour l'analyse d'imagerie médicale, illustrent l'importance de données d'entraînement de qualité pour obtenir des résultats fiables.

De surcroît, l'IA a la capacité **d'automatiser des tâches répétitives**, libérant ainsi du temps pour des interventions plus stratégiques et à plus forte valeur ajoutée. Elle s'avère également précieuse pour détecter des tendances et des corrélations qui seraient difficiles à identifier manuellement.

L'IA peut en outre fournir des traitements personnalisés adaptés aux besoins spécifiques de chaque patient.

Enfin, elle facilite l'accès aux soins en offrant des **outils de suivi et d'intervention à distance**, rendant le soutien plus accessible à tous.

Depuis l'introduction de ChatGPT, l'IA a captivé l'attention en raison de ses capacités impressionnantes et de ses nombreuses applications pratiques. Cependant, son utilisation n'est pas toujours appropriée ou optimale, ce qui soulève des interrogations quant à la manière dont elle devrait être mise en œuvre de manière éthique et efficace.

Face à la prolifération des outils l'IA et à leur intégration croissante dans le secteur de la santé mentale, il devient impératif de définir des cadres clairs pour prévenir toute dérive. C'est dans cette optique que ce document a été élaboré : **il vise à mettre en lumière les points de vigilance et à formuler des recommandations claires.**

Notre ambition est de promouvoir une utilisation consciente des nouvelles technologies parmi les professionnels de la santé mentale et les personnes concernées, patients et aidants. Nous aspirons à sensibiliser ces groupes à l'importance de mettre en œuvre ces outils de manière réfléchie, dans le respect des principes éthiques fondamentaux.

La stratégie adoptée pour intégrer les technologies émergentes, telles que l'IA, reconnaît qu'il est impossible de prédire l'ensemble des risques associés avant leur mise en œuvre à grande échelle. Ces technologies peuvent engendrer des défis et des incertitudes qui ne deviennent apparents que lors de leur déploiement et de leurs évolutions futures.

L'objectif principal est de détecter rapidement les conséquences inattendues afin d'y apporter une réponse efficace, réduisant ainsi les impacts négatifs tout en optimisant les avantages de ces innovations. Attendre la résolution de tous les risques avant d'utiliser ces technologies n'est pas envisageable, compte tenu des nombreux bénéfices significatifs que l'IA peut apporter dans divers domaines.

Afin de détecter et d'adresser ces défis de manière efficace et éthique, il est crucial d'établir un cadre national de matériovigilance spécifique à l'IA. Nous nous inspirons ici des principes de la pharmacovigilance, adaptés dans ce contexte sous l'appellation de « **numéricovigilance** ».

# Méthodologie

Ce rapport est l'aboutissement d'un processus de réflexion à plusieurs niveaux, dont les étapes sont détaillées ci-après :

1.

Dans un premier temps, nous avons entrepris de concrétiser ce projet en rassemblant diverses applications existantes, notamment celles dédiées à la détection des troubles psychiques et à l'optimisation de l'accès aux soins.

2.

Par la suite, nous avons procédé à l'identification et à la prise de contact avec des entreprises françaises et des établissements hospitaliers utilisant l'IA pour traiter ces cas d'usage, ou ayant mené une réflexion approfondie sur le sujet. Parmi les trente entreprises et hôpitaux sollicités, près de la moitié ont répondu favorablement à notre sollicitation.

3.

Chaque entretien avait pour objectif d'explorer les bénéfices de l'IA dans leur cas d'usage spécifique, d'identifier les risques inhérents à son utilisation et de débattre des solutions envisageables pour atténuer ces risques, selon les perspectives des entreprises.

4.

Pour conclure, nous avons procédé à l'analyse de ces données afin d'évaluer la probabilité et la gravité des risques identifiés, ainsi que pour élaborer de nouvelles solutions discutées au sein du groupe. Ces informations ont été consignées dans un document s'inspirant du modèle de pharmacovigilance, jugé pertinent pour évaluer l'utilisation appropriée de l'IA.





**Remarque :** Les cas d'usage et les entreprises que nous avons consultés ne prétendent pas couvrir de manière exhaustive l'ensemble des possibilités offertes par l'IA dans le domaine de la santé mentale. Il s'agit d'un état des lieux réalisé en fonction des ressources et du temps à notre disposition.

Bien que nous nous soyons efforcés d'obtenir un large éventail de perspectives, ce document se limite à refléter les informations que nous avons pu recueillir dans le cadre de notre étude. Nous nourrissons néanmoins l'espoir qu'il apportera un éclairage utile et pertinent sur les opportunités et les défis actuels liés à l'utilisation de l'IA en santé mentale.

# I. Les quatre cadres d'intervention de l'IA en santé mentale

Dans cette section et la suivante, nous proposons de segmenter les différents cas d'usage en partant de cadres d'intervention globaux (qui englobent des cas d'usage), en regroupant de manière judicieuse les résultats fournis par l'algorithme d'IA. Cette démarche permettra de renforcer l'abstraction des recommandations et de garantir que nous ne dépendons pas entièrement des algorithmes.

1.

**La prédiction de valeurs** joue un rôle crucial dans le diagnostic, où l'on exploite les données pour déterminer une valeur. À titre d'exemple, cela peut permettre d'identifier une émotion à partir d'une séquence vidéo.

2.

**La génération de texte** est utilisée pour faciliter les interactions écrites ou vocales avec les patients ou le personnel soignant. Par exemple, un agent conversationnel peut engager un dialogue en temps réel avec un patient.

3.

**La génération d'activités** est appliquée dans le cadre de jeux sérieux ou d'exercices réguliers visant à améliorer la santé mentale des patients. L'objectif est de définir un cadre pour l'activité et de la réaliser en générant du texte ou des images. À titre d'illustration, on peut concevoir un exercice pour aider un patient à réduire son niveau de stress.

4.

**La recommandation de ressources** consiste à fournir des informations pertinentes aux patients ou au personnel soignant au moment opportun. Par exemple, on peut suggérer un article sur l'anxiété à une personne cherchant à comprendre ses symptômes.

1

### PRÉDICTION DE VALEURS

Exploite les données  
pour déterminer une valeur.

Ex : identifier une émotion  
à partir d'une vidéo

2

### GÉNÉRATION DE TEXTE

Facilite les interactions écrites  
ou vocales.

Ex : Agent conversationnel pour  
dialoguer avec un patient

3

### GÉNÉRATION D'ACTIVÉS

Appliquée aux jeux sérieux et  
exercices de santé mentale

Ex : exercice pour réduire le  
stress d'un patient

4

### RECOMMANDATION DE RESSOURCES

Fournit des informations  
pertinentes au bon moment

Ex : suggérer un article sur  
l'anxiété à un patient

Les quatre cadres  
d'intervention  
de l'IA en santé  
mentale

## II. Les usages concrets de l'IA en santé mentale

Avant de détailler les différents cas d'usage de l'IA en santé mentale, il est essentiel de souligner trois axes fondamentaux qui s'appliquent à l'ensemble des applications présentées ci-après.

### ■ Intelligibilité des algorithmes

Les entreprises consultées pour ce rapport ont privilégié l'utilisation d'algorithmes intelligibles, c'est-à-dire des algorithmes dont le fonctionnement peut être compris et vérifié par leurs concepteurs. La distinction entre algorithmes explicables et non explicables pour les professionnels de santé est cruciale, car elle influence la manière dont ces outils peuvent être intégrés dans la pratique clinique. L'objectif est de maintenir la capacité des praticiens à prendre des décisions personnalisées pour leurs patients, tout en bénéficiant du support de l'IA, avec des algorithmes capables de dire "je ne sais pas".

### ■ Évaluation à long terme

Il est impératif de mener des études prospectives sur une période suffisamment longue pour évaluer les impacts réels de l'IA en santé mentale, les évaluations sur données rétrospectives mesurant la qualité de la prédiction algorithmique, mais ne présageant pas de son utilité clinique. Ces études doivent mesurer les bénéfices pour les patients et les répercussions sur les praticiens, telles que la réduction du temps de consultation, la possible délégation de certaines tâches à l'IA, ou une éventuelle perte de discernement résultant d'une confiance excessive envers l'IA. Des études comparatives avant et après l'intégration des informations fournies par l'IA sont recommandées pour quantifier ces effets.

## Protection des données

La sécurité et la confidentialité des données restent des préoccupations majeures dans l'utilisation de l'IA en santé mentale. Le risque associé aux données, intrinsèque au domaine informatique et particulièrement critique dans le secteur de la santé, expose la vie privée des patients à des violations potentielles. Ces violations peuvent entraîner des conséquences graves : un employeur pourrait décider d'écarter un collaborateur ou un assureur pourrait augmenter une franchise en se basant sur ces informations. Il est donc essentiel de mettre en place des mesures robustes pour protéger les informations sensibles des patients dans toutes les applications de l'IA.

Ces trois axes - **intelligibilité des algorithmes, évaluation à long terme, et protection des données** - constituent le cadre général dans lequel s'inscrivent tous les cas d'usage de l'IA en santé mentale présentés ci-après. Leur prise en compte est cruciale pour une intégration éthique et efficace de l'IA dans le domaine de la santé mentale.

# 1. Diagnostic et évaluation

## ■ Détection et diagnostic de troubles psychiques

Les entreprises proposent l'utilisation de l'IA pour détecter des **phénotypes numériques** associés à des troubles mentaux. Nous pouvons distinguer les « procédés passifs » qui recueillent l'information de manière non intrusive (par exemple par traitement du langage naturel ou par des capteurs vidéo) et des « procédés actifs » qui analysent la réaction à des stimuli présentés.

**Les procédés passifs** sont moins intrusifs et nécessitent généralement moins de ressources humaines. En revanche, ils utilisent beaucoup de ressources informatiques et techniques pour mettre en place l'enregistrement, la transcription et l'analyse des entretiens par exemple, pour évaluer une population cible. S'affranchir de la participation du sujet évalué permet de réduire les contraintes liées au dépistage, mais soulève la question du consentement. Il convient donc de s'assurer que le sujet évalué soit informé de manière claire, loyale et appropriée sur l'intérêt du dépistage et ses implications.

**Les dépistages actifs** utilisent des stimuli qui induisent une réaction analysée par un algorithme. L'intérêt réside dans l'augmentation de la précision des dispositifs en proposant des tests standardisés. Le risque, qui existe également pour les procédés « passifs », est celui d'une bibliothèque calibrée sur une population trop spécifique (par exemple, des images dont la connotation varierait selon la culture). L'exposition à des stimuli nécessite de s'interroger sur l'acceptabilité et l'innocuité du procédé.

Concernant le **cas d'usage d'un dépistage de troubles mentaux**<sup>(2)</sup> : le test doit être reproductible, spécifique et surtout sensible dans la population ciblée, et donc fondé sur une base de données solide, suffisamment généralisable.

Un traitement efficace du trouble doit être proposé et doit pouvoir être mis en œuvre rapidement. Un dépistage de masse des troubles psychiques implique donc de s'assurer de la disponibilité et de l'accessibilité en aval des ressources nécessaires pour traiter les troubles ainsi dépistés. Si l'intégration au sein d'une même entité des ressources pour dépister et traiter les troubles peut paraître avantageuse à cet égard, cela expose au risque d'un dépistage qui sélectionnerait une population en fonction de son éligibilité pour le traitement.

Face à cette problématique, il convient de veiller particulièrement à l'absence de conflits d'intérêts entre les sociétés impliquées dans le dépistage des troubles et celles impliquées dans leur traitement.

Nous préconisons également que le dépistage ne se réalise pas selon une logique de cloisonnement, mais s'intègre dans la logique de parcours de soins du sujet, en impliquant les structures et personnes ressources déjà engagées dans le suivi des sujets (médecin traitant, infirmier scolaire, médecine du travail, etc.)

Concernant le **cas d'usage d'une aide au diagnostic**, cela implique nécessairement la présence d'un médecin. Le risque est celui d'une mauvaise interprétation des résultats. L'annonce diagnostique étant de la responsabilité du médecin, il convient de veiller à ce que ce dernier soit formé à l'utilisation du dispositif. Nous préconisons donc une formation qualifiante pour l'utilisation des dispositifs fondés sur une analyse par une IA.

## Aide à la décision clinique

Nous retrouvons le cas d'usage d'un dispositif qui viendrait analyser les décisions cliniques des professionnels de santé. Par rapport à un système qui serait fondé sur des règles codées une par une, les algorithmes fondés sur un apprentissage automatique permettent une meilleure adaptation à des cas non vus lors de l'entraînement et une meilleure maintenabilité dans le cas de systèmes complexes. Cela se faisant en général au détriment de l'interprétabilité des résultats puisque les systèmes d'IA sont souvent des boîtes noires « magiques » : on ne sait pas toujours comment elles arrivent à leurs conclusions. Il faut aussi noter que créer ces outils d'IA demande beaucoup de temps et d'efforts, comme

(2) [https://www.has-sante.fr/jcms/c\\_2632453/fr/depistage-objectif-et-conditions](https://www.has-sante.fr/jcms/c_2632453/fr/depistage-objectif-et-conditions)

nourrir un ordinateur avec des tonnes d'informations médicales avant qu'il devienne « intelligent ».

Si le système pouvait alerter le prescripteur en cas d'erreur détectée, le gain serait également significatif en termes de iatrogénie (désigne l'ensemble des conséquences néfastes sur l'état de santé individuel ou collectif de tout acte ou mesure pratiqué ou prescrit par un professionnel de santé habilité et qui vise à préserver, améliorer ou rétablir la santé).

Les difficultés pourraient être liées à la **robustesse de l'IA** : problème lié à la base de données, en particulier des effets centres. Ceci pourrait représenter un danger si, par exemple, les systèmes d'unités utilisés dans un centre ne sont pas les mêmes que dans un autre, ou si la population d'application diffère de manière significative. L'incidence peut être importante dans le choix de préconiser une thérapeutique plutôt qu'une autre en fonction de critères financiers.

Dans cette perspective, le choix peut également être influencé par **l'accessibilité**. Par exemple, l'approche médicamenteuse ou psychothérapeutique pour traiter la dépression d'intensité modérée pourrait être choisie en fonction des possibilités d'accéder à un psychologue. En ville, les traitements peuvent être onéreux, tandis qu'ils sont disponibles gratuitement à l'hôpital, ce qui peut favoriser l'option de la thérapie par rapport à celle d'une molécule disponible à moindre coût en ville.

Nous suggérons qu'un **médecin compétent dans le domaine d'application de l'algorithme et conscient des risques mentionnés** soit impliqué dans son développement pour garantir son intégrité déontologique. De plus, nous recommandons l'établissement d'une **matérovigilance** pour vérifier l'absence de ces biais : les pratiques cliniques et les profils des patients varient souvent d'un centre à l'autre. Cela peut entraîner la présence de variables non observées et inégalement représentées entre les centres tels que le niveau socio-économique des patients, la pollution environnementale, etc. Certaines de ces variables peuvent agir comme des facteurs de confusion. Par conséquent, **prendre en compte l'effet centre permet d'ajuster les résultats en fonction de ces facteurs de confusion** (l'effet centre représente les différences systématiques entre les établissements participant à l'étude, comme les pratiques médicales spécifiques ou les équipements disponibles. En l'intégrant dans l'analyse, on peut mieux isoler l'impact réel des interventions étudiées, indépendamment des particularités de chaque centre).



Un troisième risque serait de confier la prise de décision à un algorithme. Nous serions alors en présence d'un **détournement de l'usage**, ce qui conduirait à une déshumanisation du soin. A l'heure actuelle, nous recommandons qu'une proposition algorithmique ne puisse pas être opposée au principe d'indépendance professionnelle (liberté de ne pas suivre la proposition qui va de pair avec la responsabilité du professionnel dans la prise de décision de soin). Le cadre déontologique médical est à ce titre suffisamment sécurisant. Cependant, si les performances d'un algorithme venaient à devenir significativement meilleures que celles d'un praticien, la déontologie pencherait plutôt de l'autre côté, « en premier lieu de pas nuire ».

## Évaluation des risques et pronostic

L'évaluation des risques de rechute pour en prévenir les effets constitue un défi majeur en santé mentale. Les algorithmes basés principalement sur l'apprentissage automatique sont conçus pour identifier les profils numériques à haut risque de rechute. Ces outils offrent l'avantage de repérer plus efficacement les individus à risque tout en offrant une explication potentielle plus claire que celle des cliniciens.

Le risque de ces cas d'usage est celui d'une **déshumanisation du soin** à travers un traitement statistique qui ne serait plus individualisé. Ceci est également en lien avec une **surgénéralisation par l'algorithme** qui conduirait à « enfermer » les sujets dans des catégories à risques non adaptées. La conséquence serait alors pour certains une diminution des ressources allouées pour faire face (en cas de minimisation du risque) ou au contraire l'assignation à un protocole contraignant non justifié (surestimation du risque).

La seconde catégorie de risque est celle de **l'utilisation de ces phénotypes en dehors du domaine du soin** (calcul d'une prime de risque par un assureur, évaluation du risque de burnout par un employeur...) déjà interdite par l'IA Act<sup>(3)</sup>.

Les solutions proposées par les entreprises sont la **présence d'un garant éthique humain** pour éviter le détournement de l'usage du dispositif. Un cadre déontologique d'utilisation du dispositif peut être garanti par la déontologie médicale : une solution pourrait donc être qu'un **comité pluridisciplinaire indépendant** soit nommé pour ces sociétés, qui porterait la responsabilité de l'usage éthique des algorithmes.

(3) <https://artificialintelligenceact.eu/fr/>

## 2. Traitement et suivi

### Suivi et monitoring

Le suivi et la surveillance en santé mentale s'appliquent à une diversité de patients, englobant tant ceux atteints de troubles psychiques que la population générale. À titre d'exemple, l'exploitation **d'indicateurs biométriques** tels que le niveau de stress et la fréquence cardiaque permet d'identifier précocement les manifestations physiologiques du stress et de l'épuisement par le biais de plateformes numériques de santé, notamment la téléconsultation.

L'avènement des **agents conversationnels thérapeutiques** ouvre une perspective prometteuse pour assurer un soutien motivationnel et un suivi ininterrompu aux patients. Néanmoins, il existe des risques que l'IA ne réponde pas de manière appropriée aux besoins du patient (risques d'incohérences), ce qui pourrait diminuer, voire compromettre l'efficacité du dispositif, entraînant ainsi une perte de confiance et incitant éventuellement le patient à cesser d'utiliser le dispositif.

Il incombe à l'utilisateur de signaler toute erreur qu'il détecte. Afin de minimiser ce risque, il est préconisé que le fabricant implémente un **système d'apprentissage continu** (life-long learning). Ce système permettrait de corriger les erreurs éventuelles en se fondant sur les signalements transmis automatiquement sous forme de notifications, malgré les risques de dérives côté utilisateurs.

### Thérapie et réhabilitation

#### **Réalité virtuelle (RV) : Utilisation de la réalité virtuelle pour la thérapie d'exposition et la simulation d'environnements thérapeutiques.**

La réalité virtuelle en santé mentale vise à immerger les patients dans des simulations pour les aider à faire face aux défis quotidiens, offrant ainsi des avantages tangibles pour le bien-être physique et mental. L'intégration de l'IA dans ce domaine propose plusieurs utilisations principales : fournir un soutien par le biais d'avatars amicaux ou de psychologues virtuels accessibles en continu, et proposer des environnements d'exposition permettant la désensibilisation, par exemple pour traiter certaines phobies.

Une caractéristique importante de cette technologie est son effet libérateur, permettant aux patients de s'exprimer sans craindre le jugement, favorisant ainsi le processus thérapeutique. Cependant, l'utilisation intensive de ces outils comporte des risques, notamment la **dépendance** et les **erreurs potentielles de l'IA** pouvant aggraver l'état du patient. Pour atténuer ces risques, des mesures telles que des **limites de temps d'utilisation** et des **vérifications humaines** doivent être mises en place.

### **Jeux sérieux : Développement de jeux vidéo à des fins thérapeutiques pour améliorer les compétences cognitives et émotionnelles.**

Les jeux sérieux en santé mentale peuvent revêtir différentes formes, allant des simulations interactives de situations de la vie réelle aux jeux de rôle où les joueurs peuvent explorer et résoudre des problèmes liés à leur santé mentale. Ils peuvent également inclure des éléments éducatifs, des exercices de relaxation ou de méditation, ainsi que des mécanismes de suivi des progrès du joueur.

#### **Focus Serious Gaming**

Le serious gaming, ou jeu sérieux, représente une approche innovante et prometteuse dans le domaine de la santé mentale. Ces jeux, conçus spécifiquement pour des objectifs thérapeutiques plutôt que pour le simple divertissement, offrent plusieurs avantages distincts. Ils permettent **d'engager les patients de manière interactive et immersive**, ce qui peut améliorer l'adhésion au traitement et réduire les taux d'abandon<sup>(4)(5)</sup>. Les serious games peuvent être utilisés **pour traiter une variété de troubles psychiques**, tels que la dépression, l'anxiété, le trouble de stress post-traumatique, et les troubles du spectre autistique. En intégrant des éléments de thérapies cognitivo-comportementales et d'autres techniques psychologiques, ces jeux permettent aux utilisateurs de **pratiquer et de renforcer des compétences cognitives et émotionnelles** dans un environnement sécurisé et contrôlé<sup>(6)</sup>. De plus, leur accessibilité via des plateformes numériques les rend plus facilement disponibles pour un large public, y compris ceux qui pourraient avoir des difficultés à accéder aux services de santé mentale traditionnels. En somme, le serious gaming offre une méthode complémentaire et efficace pour la gestion et le traitement des troubles psychiques, tout en rendant le processus thérapeutique plus engageant et moins stigmatisant.

(4) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8867878/>

(5) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8800090/>

(6) <https://www.collock.com/serious-game/>

Les jeux sérieux peuvent aussi servir à **diagnostiquer l'apathie**, un symptôme précurseur de la maladie d'Alzheimer. Ces outils intègrent la reconnaissance faciale pour analyser les expressions faciales et vocales, afin d'identifier d'éventuels biomarqueurs. À partir d'une série de tâches cliniques, le patient est sollicité pour lire du contenu ou s'exprimer afin d'évaluer l'apathie, comme cela a été mis en œuvre à l'Université de Berne<sup>(7)</sup> chez les patients atteints de troubles cognitifs moyens.

Le principal risque réside dans le fait que les patients utilisant cet outil peuvent éprouver de la frustration ou de la colère si l'utilisateur ne parvient pas à résoudre le jeu. La réalisation du jeu sérieux nécessite la présence d'un neuropsychologue ou d'un praticien, une **supervision humaine forte** étant impérative. Ce soutien doit être assuré par des experts compétents dans la gestion des situations de frustration et de colère.

Bien que les patients puissent être autonomes dans l'exécution du jeu sérieux, un **encadrement par un professionnel de santé** est nécessaire après environ trois ou quatre séances. Il est essentiel de fournir une formation et des conseils préventifs sur les jeux sérieux aux praticiens. De plus, il peut être bénéfique de mettre à disposition des patients une **notice spécifique**. L'outil peut être équipé de dispositifs de sécurité dans le jeu, qui détectent lorsque le patient éprouve des difficultés à progresser et lui proposent des solutions telles que : « Arrêtons pour aujourd'hui et nous reprendrons cela une autre fois ».

La présence d'une supervision humaine est cruciale, surtout en ce qui concerne les risques de dépendance, nécessitant l'accompagnement d'un praticien. Par ailleurs, les jeux sérieux pourraient évaluer et atténuer ce risque en amont, par exemple en envoyant une alerte si le temps de jeu dépasse les recommandations.

---

(7) Can a serious game-based cognitive training attenuate cognitive decline related to Alzheimer's disease? Protocol for a randomized controlled trial, Esther Brill & CO - BMC Psychiatry

## 2. Amélioration des pratiques et de l'accès aux soins

### Amélioration de l'accès aux soins

Dans le cadre de l'amélioration de l'accès aux soins, plusieurs cas d'usage ont été analysés.

L'utilisation d'**agents conversationnels** pour orienter les patients vers les ressources ou les professionnels adéquats et l'utilisation des **données géospatiales** pour optimiser la répartition des ressources et des services de santé mentale ont été étudiées.

Le risque principal repose principalement sur la **fiabilité des données collectées**. Dans le cas particulier de l'utilisation d'un agent conversationnel pour orienter les patients vers les professionnels et les ressources, le risque sur la santé des patients est identifié. En effet, une mauvaise orientation peut retarder la bonne prise en charge du patient et peut même l'en éloigner si la mauvaise orientation crée chez le patient une défiance et un renoncement à l'utilisation de l'agent conversationnel. La difficulté est de trouver le bon équilibre entre questions indispensables à l'orientation et questions relevant plus d'un diagnostic.

Enfin, les données initialement collectées par l'agent conversationnel pour faire une recommandation fiable ainsi que leur éventuel stockage sont à questionner avec soin afin d'assurer la sécurité du citoyen dans le système de santé (RGPD, réglementation propre aux Dispositifs Médicaux, etc.).

Dans les contextes de **télésurveillance**, où l'IA est employée pour améliorer la prestation des services de santé mentale à distance, une attention particulière est portée à l'exploration des outils de capture à distance de l'état émotionnel des patients.

La formation des professionnels de santé à l'utilisation de ces outils est cruciale, car elle peut avoir un impact à long terme sur leurs compétences et influencer leurs interventions en fonction des données nouvellement disponibles.

## Gestion du stress et du bien-être

Dans le cadre de la gestion du stress et du bien-être, nous avons exploré plusieurs cas d'usage, tels que l'utilisation de l'IA pour guider les utilisateurs à travers des **exercices de relaxation et de méditation**, l'utilisation de capteurs et d'IA pour **surveiller et améliorer la qualité du sommeil**, ainsi que l'utilisation de l'IA pour **suggérer des activités**, des routines ou des changements de mode de vie afin de réduire le stress, d'améliorer le sommeil, ou de faire de la prévention santé.

Dans ces situations, l'IA peut recueillir et utiliser des données provenant d'objets connectés externes, des informations fournies par l'utilisateur (telles que ses préférences, son ressenti, etc.), ainsi que des données sur l'utilisation de l'application pour formuler des recommandations. Ces recommandations forment un protocole que l'utilisateur peut choisir de suivre ou non. Il est donc important ici de mettre le plus possible l'utilisateur en responsabilité, par exemple en l'informant continuellement tout au long du protocole. **La boucle de contrôle humaine va également sécuriser la recommandation.**

Les risques résident principalement dans le protocole établi par l'IA, qui peut s'avérer inadapté pour plusieurs raisons :

- les réponses de l'utilisateur peuvent ne pas être fiables (par exemple, si l'utilisateur se ment à lui-même),
- certaines informations essentielles comme un handicap ou des pathologies diagnostiquées peuvent manquer,
- l'algorithmique, programmé pour proposer des solutions en toutes circonstances, peut faire des choix inappropriés.

**L'entraînement de l'IA est donc essentiel pour sa robustesse.** Nous pouvons imaginer que les algorithmes devraient être validés par les professionnels métiers (avec de nouveaux labels type psychothérapeute « certifié IA »), avec un seuil d'entraînement minimum.

## Formation et supervision des thérapeutes

Dans le cadre de la formation et de la supervision des thérapeutes, nous avons examiné l'utilisation de l'IA pour évaluer et fournir des retours sur les séances de thérapie, aidant ainsi à la formation continue des thérapeutes. Nous avons également discuté de l'assistance offerte aux thérapeutes dans la prise de décision et le choix des meilleures approches thérapeutiques pour chaque patient, ainsi que de l'utilisation de l'IA pour créer des scénarios virtuels destinés à la formation des thérapeutes.

Des risques ont été identifiés, tant pour le thérapeute que pour le patient.

Pour le thérapeute, la sur-utilisation de l'IA dans la prise de décision et l'évaluation comporte **le risque de déresponsabiliser le thérapeute**, potentiellement en diminuant ses compétences ou en lui ôtant le recul nécessaire à sa profession, ce qui pourrait entraîner une perte de confiance dans son intuition et son savoir-faire. De plus, cela pourrait conduire à une pratique uniforme qui ne serait pas adaptée à chaque patient.

Bien que l'utilisation de l'IA pour analyser les séances de thérapie et faciliter la formation soit bénéfique, il semble essentiel d'avoir un retour humain pour assurer un équilibre dans ce processus. La mise en place d'une boucle de contrôle humain IA-Thérapeute apprenant / Supervisé-formateur / Superviseur, est recommandée. Les données utilisées par l'IA, dont celles permettant de créer des scénarios, doivent être collectées par les métiers concernés, c'est-à-dire que **les données doivent être issues de la pratique de la thérapie par les professionnels.**

## Prévention des crises

La prédiction des crises/épisodes aigus implique l'utilisation de l'apprentissage automatique pour anticiper les risques en se basant sur des données historiques et en temps réel. Cela peut être réalisé en utilisant des indicateurs tels que la fréquence cardiaque ou même la reconnaissance faciale.

Dans le cadre de la dépression, ce scénario offrirait plusieurs avantages :

- contribuer à diminuer les coûts associés aux rechutes,
- améliorer l'efficacité du traitement,
- accroître la qualité de vie des patients en identifiant les rechutes.

Néanmoins, il convient de considérer certains risques concernant cette prédiction.

### Les « faux positifs » et « faux négatifs »

- Les « **faux positifs** » : l'IA détecte une crise alors que le patient n'en subit pas réellement. Dans ce premier scénario, il existe un risque de prescription médicamenteuse injustifiée, ainsi qu'un risque de « fatigue des alarmes<sup>(8)</sup>.»

---

(8) Phénomène où le personnel médical devient moins réactif aux alertes à force d'être exposé à de fausses alertes répétées.

- Les « **faux négatifs** » : le patient traverse une crise non détectée par l'IA. Dans cette situation, il existe un risque de retard au traitement, ce qui pourrait engendrer une détérioration de l'état du patient et potentiellement induire des comportements préjudiciables envers lui-même ou autrui.

### **La surveillance constante**

Elle concerne l'impact psychologique de la surveillance permanente sur une population vulnérable (ou même non vulnérable), qui pourrait présenter des perceptions altérées, par exemple, dans les cas d'un trouble bipolaire ou d'une paranoïa.

Pour atténuer ces risques, il est primordial de réévaluer la métrique d'évaluation de l'algorithme, car la moindre erreur peut engendrer des dommages considérables. Des mesures de précaution doivent être instaurées lors du déploiement en production, telles que des évaluations régulières avec des échelles cliniques validées, ou des consultations avec des psychiatres ou des infirmiers spécialisés.

Enfin, pour atténuer les risques psychologiques, il est recommandé de fournir un guide transparent sur l'utilisation des données et de collaborer étroitement avec les patients pour élaborer un protocole d'utilisation moins anxiogène, accompagné d'un soutien lors de l'installation du dispositif.

### **Recherche et développement en santé mentale**

Lors de cette étude, aucune entreprise engagée dans la recherche et le développement n'a répondu à nos demandes d'entretien. Nous espérons pouvoir approfondir ce point dans une version ultérieure de ce document.



# III. La réglementation en vigueur concernant l'IA peut-elle couvrir l'intégralité des risques identifiés ?

Le déploiement et l'exploitation de l'IA requièrent une vigilance particulière en matière de réglementation et de législation. Le cadre normatif européen, principalement incarné par l'IA Act et d'autres dispositions préexistantes, revêt une importance capitale dans ce contexte.

L'IA Act, ou Artificial Intelligence Act<sup>(9)</sup>, constitue un règlement de l'Union européenne visant à encadrer et à promouvoir le développement ainsi que la commercialisation des systèmes d'IA. Proposé par la Commission européenne en avril 2021, ce texte est entré en vigueur en mai 2024, à l'issue de trois années de négociations intenses. Ce règlement s'articule autour de quatre axes fondamentaux :

1.

Le respect de la réglementation de santé publique, englobant le Code de la Santé Publique, les Lois de bioéthique et le Règlement des Dispositifs Médicaux.

2.

L'observance de la réglementation spécifique à l'IA.

(9) Loi sur l'IA (2024) - Commission Européenne <https://digital-strategy.ec.europa.eu/fr/policies/regulatory-framework-ai>

3.

La protection de la vie privée, notamment par le biais du RGPD et de la Loi Informatique et Libertés.

4.

L'anticipation des risques cyber conformément à la directive NIS 2 et au Résilience Act.

L'une des évolutions majeures introduites par l'IA Act concerne la **responsabilité des produits défectueux**. Selon cette nouvelle législation, dans l'hypothèse où un produit basé sur l'IA présenterait des défauts, la responsabilité ne reposerait plus uniquement sur les contrats liant les parties concernées. Elle deviendrait **extracontractuelle**, ce qui signifie que les fabricants ou les développeurs pourraient être tenus pour responsables des préjudices causés par leurs produits et ce, même en l'absence de lien contractuel direct avec l'utilisateur final.

Dans le but d'assurer la qualité des données exploitées par les systèmes d'IA, le RGPD et l'IA Act imposent des **standards exigeants**. Cette amélioration qualitative des données facilitera également la mutualisation des espaces de données, permettant ainsi le partage et l'utilisation des données de santé de manière sécurisée et conforme aux réglementations en vigueur.

Il convient de souligner que les applications de l'IA dans le domaine de la santé, ainsi que les dispositifs médicaux numérisés, sont fréquemment classés comme présentant un niveau de risque élevé. Les réglementations relatives aux **risques des dispositifs médicaux**, bien qu'édictées en 2017, doivent désormais être conjuguées aux nouvelles exigences imposées par l'IA Act et les autres régulations. Ces efforts normatifs, quoique rigoureux, encouragent une innovation responsable et garantissent que les technologies de santé fondées sur l'IA offriront des soins sûrs et efficaces. In fine, ce cadre robuste et protecteur stimulera la confiance des patients et des professionnels de santé dans l'utilisation de l'IA.

Par conséquent, l'obtention du marquage CE, indispensable à la commercialisation des dispositifs médicaux en Europe, se complexifie avec l'avènement de l'IA Act. **Les éditeurs d'IA sont tenus de suivre six étapes cruciales afin de se conformer à ces nouvelles normes.**

1. Définir le cadre d'utilisation.
2. Obtenir et analyser le consentement des patients pour la collecte des données,
3. Évaluer l'impact global de l'IA ,
4. Anticiper les biais et les potentielles dérives de l'IA ainsi que les vulnérabilités en matière de cybersécurité,
5. Garantir la conformité réglementaire (CE, normes de l'IA Act, etc.),
6. Établir une gouvernance contractuelle appropriée.



# IV. Quelles préconisations pour atténuer les risques inhérents à l'utilisation de ces Intelligences Artificielles ?

Face aux défis et opportunités que présentent les Intelligences Artificielles dans le domaine de la santé, il est crucial d'adopter une approche réfléchie et structurée pour en maximiser les bénéfices tout en minimisant les risques.

Voici une série de recommandations visant à encadrer et optimiser l'utilisation de ces technologies innovantes.

1.

## **Élaborer une notice d'information à l'intention des utilisateurs, explicitant le fonctionnement du dispositif**

Ce document doit comporter une section dédiée à l'IA, détaillant sa finalité, ses atouts, ses risques, une description précise de la population cible et de son périmètre d'activité, ainsi que les mesures à adopter en cas de dysfonctionnement. L'information se doit d'être limpide, loyale, pertinente et adaptée à un niveau de compréhension usuel.

2.

## **Constituer un comité scientifique pluridisciplinaire**

Idéalement composé d'un médecin, un expert en IA, un éthicien et un spécialiste en réglementation, ce comité, dont les responsabilités sont clairement définies, peut être externalisé.

3.

**Impliquer pleinement le professionnel de santé dans le processus de développement algorithmique**

Bien que sa seule présence ne soit pas suffisante, ce dernier doit jouer un rôle prépondérant dans la robustesse et la fiabilité du système.

4.

**Dispenser une formation aux professionnels de santé concernant l'utilisation de l'IA et la compréhension des bénéfices qu'elle peut apporter aux patients**

Cette formation doit aborder les principes fondamentaux de l'IA, ses applications en médecine, en insistant sur l'évaluation des systèmes de machine learning. Un regard critique sur les performances de ces systèmes est indispensable.

5.

**Proposer aux utilisateurs des installations personnalisées et adaptées à leurs éventuelles appréhensions quant à l'utilisation de cette IA**

Il est impératif d'élaborer un protocole spécifique pour chaque cas d'usage. À titre d'exemple, l'utilisation de la caméra peut être rendue facultative si le patient émet des réserves quant à sa captation visuelle.

6.

**S'assurer de l'absence de conflits d'intérêts entre les entités impliquées dans le dépistage des troubles et celles engagées dans le traitement de ces mêmes affections**

7.

**Adapter les métriques d'évaluation<sup>(10)</sup> de l'algorithme en fonction du cas d'usage spécifique**

Mettre en place des mécanismes de sécurité et documenter ces protections au sein du système afin d'en garantir la transparence et l'efficacité. Il est primordial d'optimiser le score pertinent, tel que le F-score<sup>(11)</sup>. Il convient également de justifier la pertinence des métriques sélectionnées pour l'évaluation.

8.

**Retracer le cheminement logique des décisions prises par l'IA afin d'appréhender son processus décisionnel.**

Cela implique d'explicitier les raisons pour lesquelles l'IA a généré certains résultats, fondés sur ses analyses statistiques. Toutefois, une dérogation peut être envisagée si l'absence d'explicabilité est compensée par la démonstration d'un bénéfice médical substantiel et avéré.

(10) Métriques d'évaluation : Indicateurs quantitatifs mesurant les performances d'un algorithme d'IA. Exemples courants : Précision (proportion de prédictions positives correctes), Rappel (proportion de cas positifs réels correctement identifiés), Spécificité (proportion de cas négatifs réels correctement identifiés). Le choix des métriques dépend du contexte spécifique de l'application de l'IA en santé.

(11) Le F-score évalue la capacité d'un modèle de classification à prédire efficacement les individus positifs, en faisant un compromis entre la précision et le recall, à savoir quel point un système d'intelligence artificielle est bon pour repérer correctement ce qu'il cherche. dépend du contexte spécifique de l'application de l'IA en santé.

9.

**Sélectionner une population appropriée pour le contexte de l'entraînement des données et des IA, tout en demeurant réaliste.**

En cas d'impossibilité, il sera nécessaire de mener des études complémentaires afin de prévenir les biais algorithmiques. Assurer la représentativité des jeux de données d'évaluation est indispensable, faute de quoi la mesure des biais algorithmiques s'avère extrêmement difficile, voire impossible.

10.

**Comme la CNIL le recommande<sup>(12)</sup>, privilégier la parcimonie dans la collecte des informations.**

Adopter une approche pragmatique et épurée pour tester les IA, en veillant à ce que seules les données indispensables soient recueillies, sans inclure d'informations superflues.

---

(12) <https://www.cnil.fr/fr/developpement-des-systemes-dia-les-recommandations-de-la-cnil-pour-respecter-le-rgpd>

# Conclusion

L'intégration de l'IA dans le domaine de la santé marque le début d'une nouvelle ère, où la médecine se transforme profondément pour devenir plus **prédictive, personnalisée, préventive et participative**. À l'horizon 2030, nous pouvons envisager un écosystème de santé où l'IA jouera un rôle central, non seulement dans le diagnostic et le traitement, mais aussi dans la prévention et la gestion proactive de la santé.

L'étude menée par le collectif MentalTech apporte un éclairage crucial sur les opportunités et les défis liés à l'intégration de l'IA dans le domaine de la santé mentale, en examinant de près les différents cas d'usage, de la détection précoce des troubles à la formation des thérapeutes. Notre analyse a permis d'identifier les risques potentiels et de proposer des recommandations concrètes pour les atténuer. Ce travail s'inscrit bien sûr dans une démarche plus large visant à construire un écosystème de santé numérique robuste et digne de confiance, où l'innovation technologique s'allie à une réflexion éthique approfondie pour améliorer concrètement la prise en charge et le bien-être des patients.

Les algorithmes d'apprentissage profond, nourris par des données massives issues de diverses sources - dossiers médicaux électroniques, dispositifs connectés, séquençage génomique - permettront d'identifier des schémas subtils et de prédire les risques de maladies bien avant l'apparition des symptômes. Cette capacité prédictive transformera radicalement notre approche de la santé publique, **passant d'un modèle réactif à un modèle proactif**.

Parallèlement, l'IA facilitera une médecine véritablement personnalisée. Les traitements seront adaptés non seulement au profil génétique de chaque patient, mais aussi à son environnement, son mode de vie et son historique médical complet. Les assistants virtuels basés sur l'IA deviendront des compagnons de santé quotidiens, guidant les individus vers des choix de vie plus sains et facilitant une gestion continue de leur bien-être.

Cette révolution technologique en marche soulèvera par ailleurs de nouveaux défis éthiques et sociétaux. La Société devra gérer des questions complexes sur la propriété et la confidentialité des données de santé, l'équité dans l'accès aux



soins augmentés par l'IA, et le rôle évolutif des professionnels de santé. Il sera crucial de maintenir un équilibre entre l'innovation technologique et les valeurs humaines fondamentales de compassion et d'empathie dans les soins.

En définitive, l'avenir de la santé augmentée par l'IA promet des avancées extraordinaires, mais exigera une vigilance constante et une collaboration étroite entre technologues, professionnels de santé, éthiciens et décideurs politiques. C'est en relevant ces défis collectivement que nous pourrons réaliser pleinement le potentiel transformateur de l'IA pour améliorer la santé et le bien-être de l'humanité.

# Cartographie des considérations pour l'intégration de l'IA en santé

## CADRES D'INTERVENTION DE L'IA

1. Prédiction de valeurs
2. Génération de texte
3. Génération d'activités
4. Recommandation de ressource

## USAGES CONCRETS

- Diagnostic et évaluation
- Traitement et suivi
- Amélioration des pratiques
- Amélioration de l'accès aux soins

## RÉGLEMENTATION

- Respect de la réglementation de santé publique (code de la Santé Publique, Lois de bioéthique, Règlement des Dispositifs Médicaux)
- Observance de la réglementation spécifique à l'IA
- Protection de la vie privée (RGPD, Loi Informatique et Libertés)
- Anticipation des risques cyber (directive NIS 2 et Résilience Act.)

## PRÉCONISATIONS

1. Notice d'information détaillée sur l'IA pour les utilisateurs
2. Comité scientifique pluridisciplinaire (4+ membres)
3. Implication des professionnels de santé dans le développement
4. Formation des soignants à l'IA en santé
5. Installations adaptées aux appréhensions des utilisateurs
6. Prévention des conflits d'intérêts
7. Métriques d'évaluation adaptées et justifiées
8. Explicabilité du processus décisionnel de l'IA
9. Population d'entraînement appropriée ou études complémentaires
10. Collecte de données minimale (recommandation CNIL)

# Les contributeurs

## LES AUTEURS



### **Alexia ADDA**

#### **Ingénieure santé**

Directrice Générale chez KLAVA, une entreprise “Deeptech” spécialisée dans l’édition de thérapies numériques pour les troubles addictifs. Elle occupe le rôle d’Ambassadrice pour la French Care, une initiative de BPI France visant à promouvoir l’innovation en santé.



### **Arnaud Bressot**

#### **Consultant en IA spécialisé dans la santé mentale**

Améliore la pertinence et l’expérience utilisateur des applications dans ce domaine grâce à l’IA. En outre, il a créé le podcast “Minds & Machines”, qui traite des technologies au service de la santé mentale.



### **Dr Geoffrey Post**

#### **Médecin psychiatre**

Exerce une activité clinique au sein d’un cabinet de ville et du service de santé des armées. Il est également consultant en e-santé et membre du conseil départemental de l’ordre des médecins du Haut-Rhin.



### **Sabine Allouchery**

#### **Psychopraticienne**

Co-fondatrice de LISL UP, éditeur de solutions numériques de prévention de la santé mentale pour les PME. Solution globale de soutien auprès des particuliers et des salariés.

## **NOMS DES SOCIÉTÉS INTERVIEWÉES**

- **Acuity Care**
- **Callyope**
- **Dalia**
- **Agatos**
- **I-virtual**
- **My Family up**
- **Myndblue**
- **Suricog**
- **C2CARE**
- **Qare**
- **SIB**
- **SmartMacadam**

## **NOMS DES CHU ET**

## **INSTITUTS DE RECHERCHES INTERVIEWÉS**

- **CHU AIX Marseille**
- **INRIA Sophia-Antipolis**

# MentalTech

[www.mentaltech.fr](http://www.mentaltech.fr)

Octobre 2024