



**HAL**  
open science

# Le principe de garantie humaine du numérique en santé

Julie Allouche

► **To cite this version:**

Julie Allouche. Le principe de garantie humaine du numérique en santé. Ethique. 2021. dumas-03525834

**HAL Id: dumas-03525834**

**<https://dumas.ccsd.cnrs.fr/dumas-03525834v1>**

Submitted on 17 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

## **AVERTISSEMENT**

Ce mémoire est le fruit d'un travail approuvé par le jury de soutenance et réalisé dans le but d'obtenir le diplôme de Master. Ce document est mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt toute poursuite pénale.

**UNIVERSITÉ DE PARIS**

Faculté de santé

UFR Santé - Médecine

École Ethics, research, translations (ETRES)

Année 2020-2021

N° 2021REM01

**Mémoire POUR LE DIPLÔME de master 2 Éthique médicale et bioéthique**

Présenté et soutenu publiquement le 21 juin 2021

Par

**Julie ALLOUCHE**

**Le principe de garantie humaine du numérique en santé**

Dirigée par Madame le Professeur Marie-France MAMZER et Monsieur Jean-

Claude K. DUPONT

JURY

Mme le Professeur Marie-France MAMZER, PU-PH

Président

M. Marcel-Louis VIALARD, Professeur

M. Bernard ENNUYER, chercheur HDR

Mme Caroline DESPRES, chercheuse

Mme Marie MICHON, ATER en philosophie

M. Jean-Claude K. DUPONT, chercheur



*"C'est dans notre rapport à ceux qui ne représentent aucun danger que se place le véritable test éthique de l'humanité."*

- Milan Kundera

*Julie Allouche, Nuage de mots clefs d'après l'analyse des résultats*



## Remerciements

---

*À Mme le Professeur Marie-France Mamzer, ma directrice de mémoire, sans qui rien de cette aventure n'aurait pu exister. Merci pour son soutien, ses encouragements, et sa confiance.*

*À M. Jean-Claude Dupont, mon co-directeur de mémoire, pour l'accueil au sein de son laboratoire, pour sa disponibilité, sa patience et son investissement dans mon projet. Merci pour tout, sincèrement.*

*À M. David Gruson, qui fait partie de ces personnes qui font naître des vocations.*

*Aux concepteurs d'outils algorithmiques et professionnels de santé qui ont pris le temps de répondre à mes questions.*

*À mes camarades de promotion, et tout particulièrement à Lucile pour sa bienveillance et ses conseils tout au long de l'année, à Frédéric qui a partagé mon intérêt pour ce sujet dès le début de ce master, à Emma pour sa bonne humeur et les éclairages pertinents qu'elle a su m'apporter.*

*À ma famille et mes amis, qui n'ont cessé de m'encourager et de me soutenir.*

# Résumé et indexation en français et anglais

---

## **Comment le principe de garantie humaine du numérique en santé peut-il s'inscrire dans le développement des applications de l'intelligence artificielle ?**

### **Résumé :**

En matière de santé, l'introduction de l'intelligence artificielle, lorsqu'elle se matérialise par le développement d'outils d'aide au diagnostic ou à la décision thérapeutique, pose une multiplicité de questions : éthiques, sociales, juridiques, etc. L'enjeu principal est de trouver un équilibre pour d'une part, soutenir l'innovation et d'autre part, maîtriser les risques et dérives potentielles. Autrement dit : encadrer et penser notre capacité d'innovation sans la bloquer. Le principe de garantie humaine tente de répondre à cet enjeu.

L'objectif de cette étude est double : présenter le principe de garantie humaine et analyser son interaction avec l'intelligence artificielle en santé. Pour ce faire, sept entretiens semi-directifs ont été réalisés auprès de concepteurs d'outils d'intelligence artificielle et de professionnels de santé.

Cette méthode nous a permis de considérer quatre points centraux :

- sur la question de la responsabilité, où le principe de garantie humaine est insuffisant ;
- sur la question de la protection du médecin et du patient, où il est partiel ;
- sur la question de l'acceptabilité, où il pose un cadre juridique qui est une première étape pour gagner la confiance de la société ;
- sur la question de la supervision humaine, où il matérialise le rôle du médecin dans une réflexion large quant à la place à accorder à l'intelligence artificielle en santé.

Le principe de garantie humaine, en l'état actuel, est nécessaire, mais insuffisant. Comment pallier cette insuffisance ?

### **Discipline ou spécialité :**

[Sciences du Vivant [q-bio] / Éthique]

### **Mots clés français :**

[Garantie humaine / Intelligence artificielle / Supervision humaine / Chaîne de responsabilité / Acceptabilité / Protection médecin / Protection patient / Santé / principe juridique / Éthique]

# **How can the human oversight fit in with the development of artificial intelligence health applications?**

## **Abstract:**

In matters of health, the introduction of artificial intelligence, when it materializes in the development of diagnostic or therapeutic decision support tools, raises a number of questions: ethical, social, legal, etc. The main challenge is to find a balance between supporting innovation on one hand and controlling risks and abuses on the other. In other words: to frame and think our capacity for innovation without blocking it. The principle of human oversight attempts to address this challenge.

The objective of this study is twofold: to present the human oversight and to analyze its interaction with artificial intelligence in health issues. To do so, seven semi-directive interviews were conducted with designers of artificial intelligence tools and healthcare professionals.

This method allowed us to consider four central points:

- in the matter of responsibility, where the principle of human oversight is not sufficient;
- in the matter of protecting the doctor and patient, where it is partial;
- in the matter of acceptability, where it sets out a legal framework which is a first step in gaining the trust of society;
- in the matter of human supervision, where it materializes the role of the doctor in a broad reflection on the place to be given to artificial intelligence in health.

The principle of a human oversight, as it stands, is necessary but not sufficient. How can this insufficiency be addressed?

## **Discipline or specialty:**

[Life Sciences [q-bio] / Ethics]

## **English keywords:**

[Human Oversight / Artificial intelligence / Human supervision / Chain of custody / Acceptability / Physician protection / Patient protection / Health / Legal principle / Ethics]

## Liste des abréviations

---

- AN : Assemblée nationale
- AP-HP : Assistance publique – Hôpitaux de Paris
- AVC : accident vasculaire cérébral
- CCNE : Comité consultatif national d'éthique
- CNIL : Commission nationale de l'informatique et des libertés
- CSP : Code de la santé publique
- HAS : Haute Autorité de santé
- IA : intelligence artificielle
- OMS : Organisation mondiale de la santé
- RGPD : Règlement général sur la protection des données
- TDAM : traitement algorithmique de données massives



# Table des matières

---

<b>INTRODUCTION .....</b>	<b>3</b>
<b>1 : CONTEXTE .....</b>	<b>6</b>
1.1 MODÈLE FRANÇAIS DE LA BIOÉTHIQUE .....	6
1.1.1 Bioéthique en France .....	7
1.1.2 Numérique et santé .....	8
1.1.3 Développement exponentiel de traitements de données de santé .....	10
1.1.4. Essor de l'IA en santé .....	15
1.2 PRINCIPE DE GARANTIE HUMAINE ET IA EN SANTÉ .....	20
1.2.1 Proposition d'Éthik-IA .....	21
1.2.2 Textes préparatoires au projet de loi relatif à la bioéthique .....	23
1.2.3 Article 11 du projet de loi relatif à la bioéthique .....	28
1.2.4 Introduction du principe de garantie humaine à l'international .....	37
<b>2 : PROBLÉMATIQUE ET HYPOTHÈSES .....</b>	<b>39</b>
2.1 PROBLÉMATIQUE .....	39
2.2 QUESTION DE RECHERCHE ET OBJECTIFS .....	40
2.3 HYPOTHÈSES DE RECHERCHE .....	40
<b>3 : METHODOLOGIE .....</b>	<b>42</b>
3.1 CONSTITUTION D'UNE BIBLIOGRAPHIE .....	42
3.2 ÉTUDE QUALITATIVE RÉALISÉE .....	43
3.2.1 Échantillon composé de concepteurs d'outils d'IA et de professionnels de santé familiers de l'IA ...	43
3.2.2 Collecte d'informations matérialisées par des entretiens semi-directifs .....	44
3.2.3 Aspects éthiques et réglementaires .....	47
3.2.4 Méthodes et déroulement des analyses de contenu .....	48
<b>4 : RÉSULTATS .....</b>	<b>49</b>
4.1 DÉROULEMENT DE L'ÉTUDE .....	49
4.2 POPULATION DES CONCEPTEURS .....	50
4.3 RÉSULTATS THÉMATIQUES .....	50
4.3.1 Responsabilité .....	51
4.3.2 Protection médecin-patient .....	55
4.3.3 Acceptabilité du recours à l'IA en santé .....	60
4.3.4 Supervision humaine .....	63
<b>5 : DISCUSSION .....</b>	<b>66</b>
5.1 HYPOTHÈSES DE RECHERCHE ET RÉSULTATS .....	66

5.2 COMPARAISON AVEC LA LITTÉRATURE.....	72
5.2.1 Responsabilité.....	72
5.2.2 Protection médecin-patient.....	73
5.2.3 Acceptabilité du recours à l'IA en santé.....	76
5.2.4 Supervision humaine.....	77
5.3 LIMITES ET FORCES.....	77
5.3.1 Limites de l'étude.....	77
5.3.2 Points forts de l'étude.....	80
<b>CONCLUSION.....</b>	<b>82</b>
<b>BIBLIOGRAPHIE.....</b>	<b>85</b>
<b>TABLE DES FIGURES.....</b>	<b>92</b>
<b>TABLE DES TABLEAUX.....</b>	<b>93</b>
<b>ANNEXES.....</b>	<b>94</b>
<b>AUTORISATIONS DE DIFFUSION DES ŒUVRES UTILISEES.....</b>	<b>106</b>

## Introduction

---

« *Se demander ce que nous voulons pour demain, dans une perspective éthique, c'est parvenir à distinguer ce qui, à partir du possible, relève du souhaitable.* »

- CCNE, Rapport des États généraux de la bioéthique, 5 juin 2018 (1)

La protection des individus a toujours été le réel moteur de nos choix universitaires, et *a fortiori* professionnels. D'une part, par la compréhension de la pensée juridique française et aussi espagnole, en parallèle de l'apprentissage de leurs règles de droit. Il me semblait que la Justice passait par le Droit ; mais il m'a fallu très vite dissocier ces deux notions. D'autre part, par un Master 2 en éthique médicale et bioéthique, pour prendre le temps de réfléchir et de comprendre d'où venait ce besoin de protection des individus : protéger qui ? Contre quoi ? Comment ?

Si les innovations scientifiques et technologiques sont souvent synonymes de prouesses et promesses, certaines peuvent néanmoins être préoccupantes : ne risquent-elles pas de créer et/ou creuser des inégalités ? Quelles places pour le consentement et la responsabilité ? Disposons-nous du temps nécessaire pour nous adapter aux évolutions technologiques ? Comment une innovation va-t-elle modifier le type de société dans laquelle nous vivons ? Est-ce cela que nous souhaitons ? En matière de santé, toutes les innovations sont-elles réellement désirables ?

Nos recherches préliminaires étaient d'abord centrées sur les questions éthiques en santé. Ensuite, nous nous sommes intéressés au numérique pour retrouver l'aspect innovation scientifique-technologique qui nous préoccupait initialement. Enfin, nous avons essayé d'établir un lien entre ces deux notions : la santé et le numérique. C'est ainsi qu'au fil de nos recherches, nous avons établi une connexion entre la santé et le numérique grâce à la formulation d'Ethik-IA d'un nouveau principe : la garantie humaine. « *S'ouvrir résolument à l'innovation et essayer d'en réguler les enjeux au fil de son application : c'est le sens du principe d'une « garantie humaine » de l'IA en santé.* » (2)

Cette envie de travailler sur le principe de garantie humaine a été partagée par un autre camarade de classe, Frédéric Huynh, de formation médicale. Cette collaboration « médico-juridique » nous a permis de déblayer dans un premier temps, le terrain ensemble. Par la

suite, nous avons chacun étudié ce sujet sous un angle différent, selon nos formations initiales respectives tout en nourrissant des échanges réguliers afin d'enrichir nos points de vue.

Les idées reçues sur ce sujet étaient nombreuses. Il était indispensable de les identifier en amont, pour qu'elles n'influencent pas trop le travail de recherche en aval. Dans notre esprit, le *progrès* pouvait être synonyme de *danger* : la grande majorité des innovations étaient-elles mauvaises et associées à une certaine fatalité, tel que la perte de contrôle de l'Homme sur ce qu'il a créé ? Dans un futur plus ou moins proche, la machine n'allait-elle pas inévitablement dépasser l'Homme ? Enfin, pourquoi associons-nous systématiquement *innovations*, *nouveautés* à *inégalités* ?

En matière de santé, l'introduction de l'intelligence artificielle (IA) qui se matérialise par exemple, par la conception et le développement d'outils d'aide au diagnostic ou à la décision thérapeutique, pose certaines questions.

Pour le médecin, il s'agit en premier lieu de s'interroger sur la place de son autonomie décisionnelle et de sa potentielle responsabilité en cas d'erreur médicale. L'impact sur la formation des futurs médecins et enfin l'automatisation de certains emplois se posent de la même manière.

Pour le patient, ce précieux colloque singulier, précieux, car il place l'humain au cœur du soin, pourrait être bouleversé par l'arrivée d'un troisième acteur en santé : l'IA. Chaque patient est, par définition, une personne singulière avec une situation unique. Comment s'assurer que l'algorithme prendra bien en compte ces considérations alors que pour le moment il est basé sur des caractéristiques générales ? Le patient a-t-il le droit d'être informé sur le fait que son diagnostic et/ou son traitement ont été proposés initialement par une machine, notamment grâce au recueil de ses données de santé ? Comment organiser la collecte, l'organisation et la régulation des données de santé ? Le consentement du patient en matière de données de santé est-il présumé ? Dans quelle mesure l'utilisation d'un outil algorithmique utilisant de l'IA va-t-elle améliorer la qualité de la prise en charge du patient ? Le médecin et le patient pourront-ils s'opposer à l'algorithme ?

Pour le concepteur d'outils d'IA, tels que des outils d'aide au diagnostic ou à la décision thérapeutique, sa responsabilité pourrait-elle être engagée en cas de dysfonctionnement de la machine ou d'erreur médicale ? En cas d'erreur médicale et sous certaines conditions précises, serait-il possible d'envisager un partage de responsabilité entre le médecin et le concepteur ?

L'enjeu principal est donc de trouver un équilibre pour d'une part soutenir l'innovation et d'autre part maîtriser les risques et dérives, autrement dit : encadrer et penser notre capacité d'innovation sans la bloquer.

Les valeurs mises en œuvre par le développement des applications de l'IA en santé sont plurielles. Le principe d'autonomie implique d'abord le respect des personnes et de leur libre choix, ensuite de respecter la dignité des personnes et leur volonté, et enfin de traiter les patients comme des agents moraux autonomes, capables de réfléchir sur leurs objectifs personnels et d'agir conformément à cette réflexion. L'individu dispose de nombreux droits et libertés fondamentaux qui ne doivent pas être altérés par une innovation. En matière de santé, selon le principe d'équité, aucune fraction de la population ne doit subir une part excessive d'inconvénients ou être privée des avantages liés à un soin ou à une recherche.

Face à tous ces enjeux, de nombreuses solutions peuvent être envisagées.

Concernant le médecin, il paraît judicieux de renforcer son devoir d'information et garantir sa maîtrise de la décision finale. Toutefois, en pratique, ceci n'est envisageable que si le médecin comprend réellement comment l'algorithme a pu lui proposer telle ou telle proposition. David Gruson, Docteur en droit de la santé et fondateur de l'Ethik-IA, a proposé la création de « *Collèges de garantie humaine* » qui feraient office de supervision humaine en assurant la qualité des décisions rendues par les algorithmes (2).

Concernant le patient, il faudrait pouvoir l'informer de l'utilisation de traitements algorithmiques de données de santé pour assurer son consentement éclairé aux soins.

C'est pour répondre à tous ces enjeux, que Ethik-IA avait proposé en 2017, la création d'un principe de garantie humaine de l'IA en santé (3). Cette proposition avait pu être ensuite reprise dans le cadre du rapport de Cédric Villani sur l'IA (4) et, plus spécifiquement, dans celui de l'avis 129 du rapport dédié à la révision de la loi de bioéthique émis par le Comité consultatif national d'éthique (CCNE) (5). Le passage du projet de loi bioéthique au Sénat représente une nouvelle étape forte de reconnaissance et de concrétisation de ce principe général de régulation positive de l'IA en santé (6). Le principe de garantie humaine de l'IA (*Human Oversight*) a été récemment introduit au projet de règlement sur l'intelligence artificielle par la Commission européenne (7). Ce principe s'inscrit dans une logique de régulation positive visant à soutenir le développement de l'intelligence artificielle en France et en Europe, dans un cadre permettant d'en réguler les risques éthiques (8).

# 1 : CONTEXTE

---

*« Il ne peut être vain qu'une société organise des temps où est fourni l'effort de penser à plusieurs, le monde que nous souhaitons, et la façon singulière que nous avons de l'habiter. »*  
- CCNE, Rapport des États généraux de la bioéthique, 5 juin 2018 (1)

L'émergence des outils algorithmiques d'aide au diagnostic, ou d'aide à la décision utilisant l'IA pose plusieurs questions. Par exemple, en cas d'erreur médicale, les médecins et les concepteurs d'outils d'IA pourraient-ils être tenus responsables ? Ces nouvelles pratiques soulèvent des interrogations, mises en avant lors des États généraux de la bioéthique en 2018 (9). Quels sont les impacts de l'utilisation de ces outils sur la pratique des médecins et sur la prise en charge des patients ? Dans l'actuelle révision de la loi de bioéthique de 2011, il est question, à l'article 11, de la création d'un nouveau cadre juridique pour faire face aux enjeux de l'utilisation de l'IA en santé : le principe de « garantie humaine » (6). Ce principe est introduit à l'échelle européenne depuis avril dernier à l'article 14 du projet de règlement de l'intelligence artificielle (7).

Il semble nécessaire de comprendre à quels besoins répond le principe de garantie humaine. Pourquoi et comment en est-on venu à proposer sa création ? À quoi correspond-il concrètement ? La conception française est-elle similaire à la conception européenne du principe de garantie humaine ?

L'objet de cette partie est de dresser un panorama complet du principe de garantie humaine à partir de la littérature existante à ce sujet. Pour comprendre dans quel contexte a émergé ce principe, nous ferons un pas en arrière et étudierons les spécificités du modèle français de la bioéthique.

## 1.1 MODÈLE FRANÇAIS DE LA BIOÉTHIQUE

Le développement massif des données de santé et des applications l'IA en santé est venu bousculer les pratiques médicales et générer des inquiétudes sociétales. Ces inquiétudes ont

été mises en avant lors des États généraux de la bioéthique de janvier à avril 2018. L'objectif des États généraux de la bioéthique était de « *permettre aux citoyens de s'approprier, par une consultation numérique ou lors de débats locaux, les thématiques à l'ordre du jour de la révision* [de la loi de bioéthique] » (5).

D'une part, nous analyserons la conception française de la bioéthique et le lien qui unit le numérique à la santé. D'autre part, ce constat préalable nous permettra de faire un état des lieux du développement des données de santé et de l'IA en santé.

L'objet de cette sous-partie est de voir comment le développement des données de santé et l'essor de l'IA en santé bouleversent le traditionnel modèle français de la bioéthique et remettent en question nos pratiques.

### **1.1.1 Bioéthique en France**

« *Toute révision de lois de bioéthique relève de l'art du mikado* » (10). Cette métaphore du Conseil d'État illustre la difficulté d'apporter une modification ponctuelle sans ébranler l'édifice dans son ensemble.

Tracer le périmètre de la réflexion bioéthique est un exercice délicat. Dès 1927, le Pasteur allemand Fritz Jahr a forgé le terme de « bio-ethik » (11), mais l'expression ne s'est imposée qu'à la suite de la publication en 1971 du livre de Van Renssaler Potter *Bioethics : bridge to the future*. Ce dernier définit la bioéthique comme une combinaison des connaissances biologiques (*bios*) et des valeurs humaines (*ethos*) et se propose d'examiner l'interrelation entre la morale et les sciences de la vie. L'approche a ensuite été circonscrite aux difficultés posées par le développement de nouvelles pratiques médicales, compte tenu des préoccupations immédiates liées à l'apparition des techniques de reproduction artificielle (12).

Dans le domaine de la bioéthique, la France a été pionnière en matière de législation. La première, dite loi Huriot-Sérusclat, remonte à la fin des années 1980 (13). Le principal thème abordé était la protection de la personne dans le cadre des essais cliniques (recherche biomédicale). Aujourd'hui, la loi de bioéthique en vigueur date de 2011 (14). Elle aborde différents thèmes, tels que l'interruption médicale de grossesse ou le don d'organes, mais n'aborde pas les données de santé et l'IA. Elle fait actuellement l'objet d'une révision au

Parlement. Le récapitulatif historique des lois de bioéthique en France proposé par le CCNE est disponible en annexe 1 (1).

Le modèle français de bioéthique repose sur la conciliation de trois principes : la dignité, la liberté et la solidarité. Le principe du respect de la dignité de la personne humaine impose que « *la personne ne soit jamais considérée seulement comme un moyen, mais aussi une fin, qu'elle ne soit pas instrumentalisée* » (15). La conception française particulièrement protectrice du corps humain a conduit le Conseil constitutionnel à conférer une valeur constitutionnelle au principe de dignité de la personne humaine. (16) L'attention accordée à l'information et au consentement du patient, et le respect de son autonomie dans ses choix témoignent quant à eux de l'importance reconnue au principe de liberté (10).

Pour veiller à la sauvegarde du modèle français de la bioéthique lors de la révision des lois de bioéthique, le Premier ministre, Édouard Philippe, a écrit au Conseil d'État, le 6 décembre 2017. Dans cette lettre, il souhaite s'assurer que les différents enjeux juridiques liés aux questions à l'ordre du jour de la révision soient toujours remis en perspective des principes de dignité, de liberté et de solidarité. Cette lettre est reproduite en annexe 2.

Toutefois, ce modèle français est mis en tension par plusieurs évolutions :

- « *L'essor de l'IA et le développement exponentiel de traitements de données de santé [...] suscitent des questions nouvelles, porteuses d'espoirs médicaux sans précédent, mais également de menaces pour la protection des personnes et de leur vie privée et même pour l'évolution de l'espèce humaine* » (17).
- Les aspirations de la société française témoignent d'un assouplissement des règles fixées par les précédentes lois de bioéthique aux noms des principes de liberté et d'autonomie. C'est d'ailleurs pour cela que le débat autour de la gestation pour autrui est possible (10).
- Le modèle français entre en concurrence avec d'autres modèles étrangers. Ce phénomène conduit parfois le législateur à se positionner sur des sujets, car il ne peut plus continuer à contourner ses enjeux. Par exemple, si la gestation pour autrui est interdite en France (18), mais qu'elle est autorisée au Royaume-Uni, la question de la reconnaissance de la filiation en France de ces enfants est dès lors inévitable.

### **1.1.2 Numérique et santé**



*« Il existe un écart entre ce qui est techniquement possible et ce qui est éthiquement souhaitable, un écart qui légitime la réflexion éthique. »*

– CCNE, *Avis 129* (15)

#### 1.1.2.1 Définitions

Le **numérique** *« concerne des nombres, se présente sous la forme de nombres ou de chiffres, ou concerne des opérations sur des nombres »* (19).

L'Organisation mondiale de la santé (OMS) définit la **santé** comme *« un état de complet bien-être physique, mental et social, et ne consiste pas seulement en une absence de maladie ou d'infirmité »* (20).

Quel lien peut-on établir entre le numérique et la santé ? Les avancées techniques du numérique ont largement impacté le domaine de la santé, avec par exemple l'invention de l'imagerie médicale. Cependant le CCNE s'interroge sur ces notions : *« on ne peut considérer toute avancée technique comme un progrès : certaines avancées techniques peuvent contribuer à dégrader la qualité de la vie et de la santé d'une partie de l'humanité et à avoir des conséquences regrettables pour l'humanité. Le possible n'est pas toujours souhaitable. »* (15).

Pour comprendre les enjeux autour de la construction du principe de garantie humaine, il est important, à titre préliminaire, de bien faire la distinction entre le développement de trois types d'algorithmes qui correspondent à des moments différents du parcours de soin du patient :

- outil d'aide au diagnostic.
- outil d'aide à la décision médicale et thérapeutique (21)(22).
- outil d'aide à la prise en charge de la prescription médicale (23).

Un déficit de compréhension a été souligné lors des États généraux de la bioéthique. Il se traduit par une demande très forte d'information dans les domaines où l'innovation technique a été la plus rapide et novatrice (5). Ce déficit peut sembler paradoxal à l'heure de la généralisation de l'accès à Internet, mais illustre que ce moyen de communication de l'information scientifique peut donner *« l'illusion d'une connaissance universelle accessible à tous »* alors que le niveau de connaissance est insuffisant (24).

### 1.1.2.2 Enjeux éthiques

Les enjeux éthiques du développement des algorithmes au service de la médecine sont nombreux. Voici deux risques majeurs en lien avec notre sujet, qui ont été mis en avant dans l'avis 129 du CCNE (15) :

- Priver le patient, face aux propositions de décisions fournies par des algorithmes, d'une large partie de sa capacité de participation à la construction de son processus de prise en charge.
- Priver le patient de la prise en compte de ses spécificités et *a fortiori* lui proposer un traitement inadapté. L'outil fonctionnerait seulement sur la population sur laquelle il a été entraîné et ne prendrait pas en compte les caractéristiques individuelles de chaque personne. Il ne serait pas conçu pour fonctionner sur une population générale. Or les patients ne réagissent pas tous de la même façon aux traitements. Le risque est donc d'avoir des « laissés pour compte ».

### 1.1.3 Développement exponentiel de traitements de données de santé

#### 1.1.3.1 Définitions

« *Data drives all we do.* » (« *Les données déterminent tout ce que nous faisons.* »), tel est le slogan de l'entreprise Cambridge Analytica. Cette entreprise se donne pour mission de « *changer le comportement grâce aux données* » et fonctionne en mélangeant le traitement quantitatif de données, la psychométrie et la psychologie comportementale (25).

Une **donnée** est une construction de la réalité. Le Règlement général sur la protection des données (RGPD) définit à son article 4, une **donnée à caractère personnel** comme « *toute information se rapportant à une personne physique identifiée ou identifiable directement ou indirectement* » (26).

Dans ce même article, une définition juridique des **données de santé** est introduite : « *Les données à caractère personnel concernant la santé sont les données relatives à la santé* »

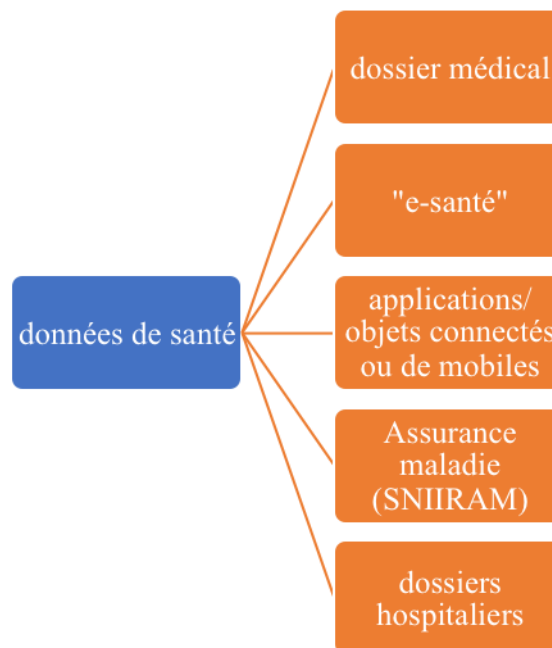
*physique ou mentale, passée, présente ou future, d'une personne physique (y compris la prestation de services de soins de santé) qui révèlent des informations sur l'état de santé de cette personne. » (26).*

La Commission nationale de l'informatique et des libertés (CNIL) a tiré trois critères de cette définition pour qualifier une donnée de santé :

- 1) La nature même de la donnée de santé ;
- 2) L'apparition d'informations relatives à la santé du fait du croisement de données ;
- 3) La finalité d'utilisation de la donnée.

Aussi, la CNIL souligne que « *cette notion recouvre non seulement l'ensemble des données collectées et produites dans le cadre du parcours de soins, mais aussi celles qui, détenues par d'autres acteurs (développeurs d'application par exemple), constituent une information sur l'état de santé de la personne* » (27).

Figure 1 : Sources des données de santé



Source : Julie Allouche, d'après : CCNE, *Rapport des États généraux de la bioéthique* du 5 juin 2018, 11/05/2021

### 1.1.3.2 Droit positif

En France, on opère une distinction entre les données personnelles relatives à la santé, collectées dans le cadre du soin ou collectées hors du cadre institutionnel du soin. Cette distinction a un impact important sur le régime juridique applicable.

Tableau 1 : Régimes juridiques applicables aux données de santé

<b>Collectées DANS le cadre du soin</b>	<b>Collectées HORS du cadre du soin</b>
Loi n°78-17 du 06/01/1978 dite informatique et liberté (28) <ul style="list-style-type: none"> <li>➤ Fondements de la protection des données.</li> <li>➤ Création de la CNIL qui a un rôle de contrôle de traitement des données.</li> </ul>	Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 (26)
Loi n°2002-303 du 04/03/02 relative aux droits des malades et à la qualité du système de santé.(29) <ul style="list-style-type: none"> <li>➤ Démocratisation du champ sanitaire.</li> </ul>	Aucune protection particulière en droit français !
Loi n°2016-1321 du 07/09/16 pour une République numérique (30) <ul style="list-style-type: none"> <li>➤ Article 49 relatif au principe de loyauté. .</li> </ul>	
Loi n° 2018-493 du 20/06/18 relative à la protection des données personnelles (31)	
Article L. 1110-4 du Code de la santé publique (CSP) relatif au secret médical et aux droits des personnes malades et des usagers du système de santé (32)	

Source : Julie Allouche, d'après : CCNE, *Rapport des États généraux de la bioéthique* du 5 juin 2018, 11/05/2021

Il est donc important de bien distinguer parmi les données de santé, celles qui sont récoltées dans le cadre institutionnel du soin, par exemple le dossier d'un patient ; de celles qui le ne le sont pas, car le cadre législatif n'est pas le même. Par exemple, les montres connectées collectent des données de santé hors cadre institutionnel, telles que la mesure du

rythme cardiaque. Ces données ne sont pas considérées comme sensibles et n'auront pas le même traitement. Le régime juridique des données de santé collectées dans le cadre du soin est plus protecteur que celui des données de santé collectées en dehors du cadre du soin.

### 1.1.3.3 Enjeux éthiques

*« La notion même de « personne » ne se limite plus à son corps, mais se démultiplie dans les données numériques de santé la concernant, qui sont échangées, stockées, marchandisées et échappent largement à son contrôle. »*

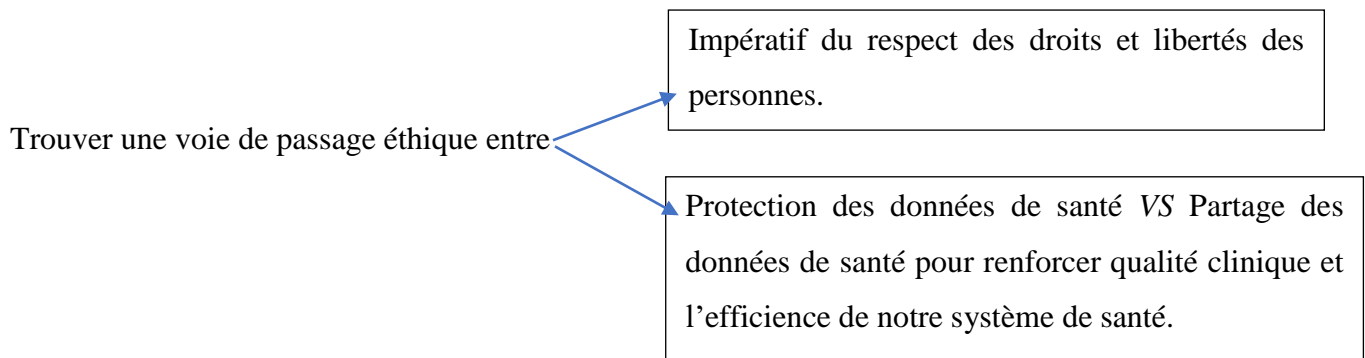
- CCNE, Avis 129, 25 septembre 2018 (15)

Les enjeux éthiques qui découlent des risques potentiels autour de la récolte et de l'utilisation des données de santé sont nombreux :

- Les points de vue des patients et des professionnels de santé vont permettre de produire et de co-construire des données. L'interaction structure la qualité et la quantité des données. Or, aujourd'hui, c'est l'ordinateur qui détermine l'organisation de la consultation. L'entretien risque de ne pas être fluide et cela peut verrouiller le discours du patient. En ce sens, le développement de la télémédecine pourrait détériorer ou du moins modifier la relation médecin-patient et la prise en charge, car *« comment percevoir l'empathie de mon médecin par la télémédecine ? »* (33). Il existe un risque qu'à terme, la décision médicale soit imposée par l'outil numérique et non plus explicitée et partagée entre le médecin et le patient. Il est également à craindre que la télémédecine et les objets connectés constituent la seule option proposée pour pallier les insuffisances du système de santé. Comment veiller à ce que l'exploitation des données ne se substitue pas à la décision humaine dans le soin ?
  
- Les données ne sont pas représentatives de toute la population, car elles ne prennent pas en compte certains groupes de la société comme les personnes en situation de précarité ou de fragilité. De plus, il peut manquer des données dans le dossier médical, comme par exemple, les antécédents médicaux de personnes qui ont perdu les liens familiaux ; ou encore ce qui n'a pas été noté dans celui-ci.

- La part de subjectivité est difficile à faire ressortir. La machine peut reproduire les biais humains. Les **biais** sont des déviations systématiques de la pensée logique et rationnelle, par rapport à la réalité. Par exemple, Amazon s'est rendu compte que son outil de recrutement avait un biais sur les femmes, car il excluait systématiquement les *curriculum vitae* des femmes. L'outil numérique reproduit ainsi les erreurs humaines (34).
- L'anonymisation n'est-elle pas utopique, connaissant la réalité des possibilités actuelles de désanonymisation des données ? En effet, toute distinction des données selon quelles concernent ou non la santé, ne résiste pas au croisement des données par les algorithmes. On pourrait tout à fait s'interroger sur le sens même de l'expression « données de santé » ...
- Le consentement individuel est affecté par la récolte et l'utilisation des données puisqu'elles ne concernent pas que la personne, mais aussi ses descendants. Le droit de chacun de les céder librement pourrait être limité. Quel pourrait être le devenir des données de santé des patients mineurs ? La liberté individuelle ne risque-t-elle pas de s'effacer face à la généralisation du numérique ? (1)
- La collecte de données peut également être un problème ; certains systèmes informatiques peuvent ne pas résister aux cyberattaques, comme le témoignent les récents piratages des systèmes informatiques hospitaliers. L'Assistance Publique – Hôpitaux de Paris (AP-HP) a été victime d'une cyberattaque alors que « *des centaines de patients atteints par la Covid-19 affluaient dans les services de réanimation débordés. [ ... ] Pendant une heure, l'AP-HP a dû couper l'accès aux e-mails et aux outils à distance pour ses salariés en télétravail* » (35). Sommes-nous capables de travailler sans la machine ? Comment assurer la protection des patients tout en permettant les progrès médicaux ?
- L'exploitation des données pourrait engendrer un contrôle social et un risque de profilage. À titre d'illustration, le taux de remboursement pourrait être lié à l'observance médicamenteuse. Comment assurer la protection du secret médical et la confidentialité des données ?

Figure 2 : Enjeux éthiques du numérique et santé



Source : Julie Allouche, d'après : CCNE, Avis 129 du 25 septembre 2018, 11/05/2021

#### 1.1.4. Essor de l'IA en santé

*« Les progrès de la médecine lui font courir un danger auquel elle ne va pas peut-être plus pouvoir résister : celui d'oublier, à côté de ses humeurs, l'homme qui est son objet, l'homme total, être de chair et de sentiment. Et d'instinct, elle se demande s'il ne faudrait que soit remise à sa place l'éminente primauté de l'observation de l'homme par l'homme, afin qu'on ne voie pas s'effondrer le vieux sens hippocratique devant la dictature des appareils ».*

– René Leriche, *Leçon inaugurale au Collège de France*, 1938 (36)

##### 1.1.4.1 Définitions

L'**intelligence artificielle (IA)** peut se définir comme telle : *« créé en 1956 par l'Américain John McCarthy, le terme « intelligence artificielle » est depuis passé dans le langage courant. Selon Marvin Minsky, un autre pionnier de l'intelligence artificielle, elle correspond à la « construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains, car elles demandent*

*des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique ». Né formellement lors de la conférence de Darmouth en 1956, le concept rencontre depuis quelques années un renouveau d'intérêt. » (1, p.96)*

Un **algorithme** pourrait se définir comme tel : « *au sens strict, description d'une suite finie et non ambiguë d'étapes (ou d'instructions) permettant d'obtenir un résultat à partir d'éléments fournis en entrée. Dans le monde numérique actuel, les algorithmes informatiques permettent de combiner des informations les plus diverses pour produire une grande variété de résultats: simuler l'évolution de la propagation de la grippe en hiver, recommander des livres à des clients sur la base des choix déjà effectués par d'autres clients, comparer des images numériques de visages ou d'empreintes digitales... » (1, p.110)*

Et enfin, le terme de **machine learning** (machine autoapprenante) se définit ainsi : « *il s'agit d'alimenter en données des algorithmes d'apprentissage, afin que les machines puissent déterminer d'elles-mêmes, sur la base d'exemples qui lui sont soumis, les opérations à effectuer pour accomplir les tâches qui lui sont assignées. Les résultats des algorithmes dépendent majoritairement des données fournies » (10).*

À titre liminaire, il convient de souligner que, tous les algorithmes n'utilisent pas de l'IA. De même, toutes les IA ne sont pas des machines autoapprenantes.

#### 2.1.4.2 Droit positif

Les législations européennes et internes relatives aux données de santé s'appliquent en matière d'intelligence artificielle, exceptée la loi relative aux droits des malades et l'article L. 1110-4 du Code de la santé publique (*cf. supra* la partie sur les données de santé, sous-section droit positif en France).

Dans le régime juridique actuel, si l'IA relève du droit des biens, les dommages susceptibles d'être causés par l'IA relèvent quant à eux de la responsabilité du fait des choses. Le Conseil d'État recommande de distinguer deux situations concernant la responsabilité du médecin dans l'usage d'outils algorithmiques intégrant de l'IA :



1. « **En l'absence de défaut établi du dispositif médical**, le médecin pourrait voir sa responsabilité engagée pour faute. Une telle faute ne saurait être établie au seul motif que le praticien n'a pas suivi les prévisions du système d'IA, même dans l'hypothèse où ces dernières se révéleraient exactes ».
2. « **Si le défaut du dispositif médical utilisé est établi**, la situation est plus compliquée, car on pourrait envisager à la fois un régime de responsabilité du fait des produits défectueux (particularité des machines autoapprenantes) tout en mobilisant différents régimes de responsabilités existants ». (10)

La question d'élaborer de nouveaux régimes juridiques adaptés aux applications nouvelles de l'IA en santé s'est posée. En effet, les régimes de responsabilité actuels n'ont pas été pensés pour répondre à l'utilisation des outils algorithmiques utilisant l'IA. Toutefois, le Conseil d'État s'est prononcé en faveur d'un maintien des régimes de responsabilité actuels. Il souhaite d'abord voir comment vont évoluer les progrès techniques et les orientations des systèmes assurantiels. En ce sens, il s'est prononcé en défaveur de l'attribution d'une personnalité juridique aux dispositifs d'IA. Pour le moment, leur fonctionnement dépend de paramétrages effectués par des êtres humains (10).

Pour la CNIL, la régulation de l'IA repose sur deux principes fondateurs pour garantir le consentement libre et éclairé du patient.

Figure 3 : Les principes de régulation de l'IA



Source : Julie Allouche, d'après : CCNE, *Rapport des États généraux de la bioéthique* du 5 juin 2018, 11/05/2021

1. Le principe de loyauté : « un algorithme ne doit pas entrer en opposition avec les intérêts des utilisateurs et des citoyens. Il doit offrir une certaine transparence. » (37).

2. Le principe d'obligation de vigilance : « *permettre l'organisation d'une forme de questionnement régulier, méthodique et délibératif à l'égard de ces traitements algorithmiques* » (37).

Appliqué au domaine de la santé, il s'agit de permettre à l'Homme de garder la main, plus encore que sur sa vie privée, sur son intimité. La CNIL rappelle qu'il est nécessaire de personnaliser les outils d'IA en fonction des besoins et de ne pas nourrir une confiance excessive en eux.

#### 2.1.4.3 Enjeux éthiques

- La « boîte noire » : dans certains cas, la machine continue de fonctionner sans aucune transparence. Les résultats sont ininterprétables, car excessivement complexes et obscurs pour l'humain, on ne peut donc que la croire, sans pouvoir expliquer ses choix. On sait que le dispositif fonctionne, mais on ne sait pas comment. Le problème se pose surtout en santé, car le médecin ne serait pas capable d'expliquer le choix de la machine au patient, car lui-même ne le comprendrait pas (4).
- L'exigence d'explicabilité et d'interprétabilité des dispositifs d'IA : si le professionnel de santé ne comprend pas comment le dispositif d'IA en est venu à lui proposer tel ou tel diagnostic pour son patient, il risque de ne pas être capable de s'en écarter si besoin (7). Si l'exigence d'explicabilité est désirée dans tous les domaines, elle est nécessaire en médecine, car le médecin engage sa responsabilité en cas d'erreur médicale. Il serait souhaitable que les concepteurs d'outils d'aide à la décision utilisant de l'IA veillent à ce que les résultats soient autant que possible interprétables et explicables par les utilisateurs (1).
- La responsabilité : Lorsqu'une décision médicale est erronée et que celle-ci a été prise sur la base d'un outil algorithmique intégrant de l'IA, il convient de vérifier si un défaut du dispositif médical, par exemple un dysfonctionnement de la machine ou des données biaisées, pourrait être à l'origine de cette erreur. Pour les internautes participants aux États généraux de la bioéthique, le principe d'une personnalité électronique du robot est fermement rejeté, l'humain ne pouvant se décharger de ses responsabilités sur un robot (1). En cas d'erreur dans le diagnostic médical ou dans le traitement prescrit, qui sera considéré comme responsable ? Le médecin ? Le

concepteur de l'algorithme ? Le programmeur ? Le gestionnaire de base de données ayant servi à l'apprentissage ? L'enjeu principal est de garantir qu'un patient victime de l'erreur d'un algorithme puisse être indemnisé de son préjudice.

- « L'illusion de relation humaine » (1) : le développement des applications de l'IA en santé pourrait avoir un impact sur le lien humain et sur la relation médecin-patient. Le médecin pourrait se sentir exclu. La machine pourrait relativiser sa compétence médicale et altérer la confiance du patient plus enclin à croire la machine que son médecin. À l'inverse, le patient pourrait lui aussi se sentir exclu. Il pourrait être privé de la possibilité de participer aux décisions et choix thérapeutiques le concernant. À juste titre, l'Assemblée nationale rappelle que les algorithmes ne peuvent être qu'un complément à l'activité du médecin et en aucun cas le substituer (38). Continuer à maintenir la confiance dans la relation médecin-patient est donc primordial. Une étude menée par M.-C. Lai, M. Brian et M.-F. Mamzer (39) s'intéresse d'une part, à la façon dont les professionnels de santé français perçoivent l'arrivée de l'IA dans leur pratique quotidienne ; d'autre part, à la perception des autres acteurs concernés par l'IA (les industriels de santé, les acteurs institutionnels et les chercheurs en santé spécialisés dans l'IA). L'objectif de cette étude est de comprendre en quoi le discours tenu en public sur l'IA en santé n'est pas toujours en corrélation avec son impact réel et les conséquences de cet écart.
  
- Diagnostic sans intervention humaine : Il est essentiel de veiller à ce que la décision médicale ne soit pas imposée par l'outil numérique, au risque de ne plus être explicitée et partagée entre le médecin et le patient. Pour cela, il faudrait « consacrer explicitement l'interdiction d'un diagnostic établi uniquement par un système d'IA sans intervention d'un médecin. » (10).
  
- Cristallisation du savoir : Si l'IA prend les décisions, quelles sont les données d'entraînement ? Comment construire l'expertise ? Le développement des outils d'aide au diagnostic, exemple pour dépister les naevus malins, fait que le savoir expérientiel et la formation des médecins par la clinique diminuent. Il se pourrait que la médecine clinique soit de moins bonne qualité à cause d'un recours massif à ces outils. Comment créer de nouvelles connaissances ? Qui a accès à ces données ? Nous pouvons par exemple citer le scandale Cambridge Analytica : l'entreprise Facebook

collecte une masse considérable de données personnelles. Elle n'a pas ouvert ces données aux sociologues. Pourtant, elle a accepté que l'entreprise britannique Cambridge Analytica analyse les données de dizaines de millions d'utilisateurs à leur insu. Donald Trump a d'ailleurs eu recours à cette entreprise durant sa campagne pour la présidentielle de 2016 aux États-Unis. Le vote de millions d'individus a pu ainsi être influencé en modifiant leur comportement grâce aux données (25).

- L'accessibilité à ces nouvelles technologies : Comment s'assurer que ces dispositifs d'IA soient déployés sur l'ensemble du territoire ? L'objectif est de ne pas exacerber l'inégalité territoriale d'accès aux soins. Ainsi, une grande partie de la population française se situe « *hors du numérique en général, soit par méconnaissance, soit par impossibilité d'accès* » (40). Le principe d'équité dans l'accès aux meilleurs soins pourrait être bafoué si toute la population n'a pas accès à ces nouvelles technologies.

L'étude du modèle français de la bioéthique a permis de mettre en avant deux constats : si l'IA a des potentialités majeures pour l'amélioration de la qualité du système de santé, son développement s'accompagne de nombreux risques éthiques. Dès lors, comment le principe de garantie humaine du numérique en santé peut-il s'inscrire dans le développement des applications de l'IA en santé ?

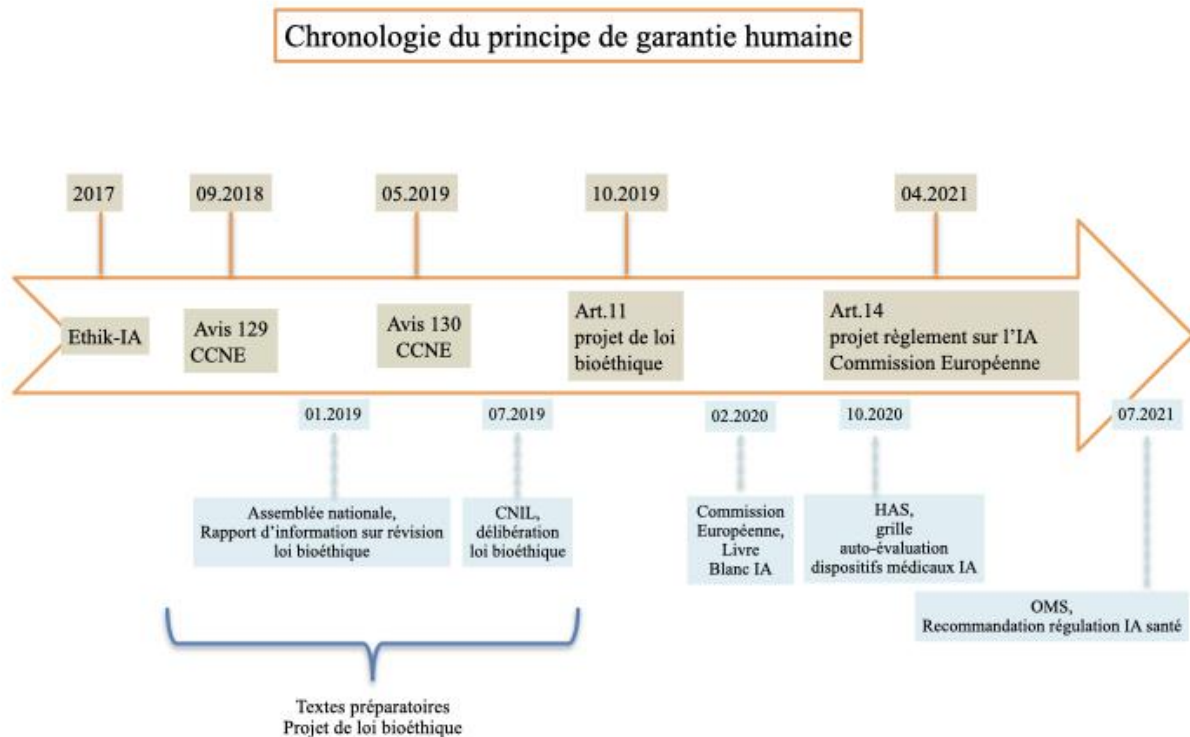
## **1.2 PRINCIPE DE GARANTIE HUMAINE ET IA EN SANTÉ**

*« Ce principe de garantie humaine, au fond c'est un peu comme la démocratie : c'est un peu le pire des systèmes, mais après tous les autres. »*

- David Gruson

L'étude du principe de garantie humaine se déroulera en trois temps : d'abord, nous étudierons la genèse et la portée de ce principe, précédemment à son introduction à l'article 11 du projet de loi bioéthique. Ensuite, nous analyserons l'évolution de ce principe dans les versions successives de l'article 11 proposé par l'Assemblée nationale et le Sénat. Enfin, nous aborderons l'introduction du principe de garantie humaine à l'international.

Figure 4 : Chronologie du principe de garantie humaine



Source : Julie Allouche, 07/06/2021

### 1.2.1 Proposition d’Ethik-IA

La rédaction de cette sous-partie a été permise par différentes sources : un échange par visioconférence en mai 2021 avec Monsieur David Gruson, deux fiches thématiques sur le principe de garantie humaine d’Ethik-IA (8) (41), notre participation au webinar organisé par CARE INSIGHT le 25 mai 2021 intitulé « *Conformité en garantie humaine de l’IA en santé : Comment s’y préparer ?* » (42).

La notion de garantie humaine provient tout d’abord de la fiction. En 2018, David Gruson publie un polar intitulé S.A.R.R.A. qui décrit une situation de risque épidémique où une IA, S.A.R.R.A., produit des vaccins (43). Pour David Gruson, on ne pouvait pas rendre crédible une IA qui produit des vaccins, sans imaginer une supervision humaine. C’est à

l'occasion de la relecture de son livre, que certaines personnes de son entourage lui ont fait prendre conscience de l'intérêt juridique de ce principe !

Ethik-IA est une initiative académique et citoyenne lancée en 2017. Fondée par David Gruson, elle réunit une équipe d'enseignants, de chercheurs en droit numérique, en technologie de l'information et de la communication et en sciences humaines et sociales. Cette initiative entend défendre une régulation positive de l'IA et de la robotisation en santé.

En 2017, elle introduit le principe de garantie humaine de l'IA en santé. Cette proposition émane de deux préoccupations principales (44):

1. L'ouverture à la technologie : le retard numérique en France impose de s'ouvrir à la technologie pour améliorer la qualité du système de santé. Il ne serait pas éthique de bloquer la diffusion de l'IA, car celle-ci peut être un outil de qualité pour les médecins et pour la prise en charge des patients. Paradoxalement, une sur réglementation de ces technologies, pour se protéger des risques fantasmés de l'IA, favoriserait l'importation de technologies issues de cadres juridiques extraeuropéens moins protecteurs que le cadre juridique français.
2. Les risques éthiques associés à l'IA en santé sont de trois ordres :
  - 2.1. Risque sur les régimes de responsabilité juridique des professionnels de santé : pour l'instant, les régimes de responsabilité du fait des choses et des produits défectueux suffisent à aborder le risque associé aux dommages de la machine. Toutefois, cette préoccupation occupera une place grandissante dans les prochaines années.
  - 2.2. Risque de délégations : ce risque est double. Du côté du médecin, le risque de délégation de la décision médicale est favorisé par la confiance placée par le médecin dans l'outil d'IA. Par exemple, si l'IA propose une solution au médecin qui s'appuie sur 500 000 cas avec 99 % de recommandations, le risque est que le médecin ne se pose plus de questions, et qu'il perde ainsi son recul critique. Du côté du patient, le risque de délégation du consentement suit un raisonnement similaire. Le patient risque d'accepter systématiquement la solution proposée par la machine, sans avoir de recul.

2.3. Risque de minoration de l'intérêt individuel VS l'intérêt collectif : dans le monde algorithmique, l'IA est fondée sur la loi du plus grand nombre. Le risque est de ne plus prendre en compte les singularités, les minorités sous représentées qui pourraient être ainsi noyées dans la masse du plus grand nombre.

Selon David Gruson, le sens du principe de garantie humaine est de trouver « *une voie de passage, par essence pas absolument satisfaisante, entre ces deux polarités : s'ouvrir à la technologie, pour éviter de se la faire imposer venue d'ailleurs, et réguler les risques de délégations et les risques de minorations de l'intérêt de l'individu versus le collectif* » (44).

La garantie humaine a été calibrée selon deux paramètres par Ethik-IA (8) :

1. Une obligation d'information du patient sur le recours à l'IA dans sa prise en charge : pour prévenir le risque de délégation de consentement du patient. Cette information est préalable à l'application de l'algorithme à sa prise en charge.
2. Une supervision humaine : à deux niveaux stratégiques pour ne pas bloquer l'innovation. D'une part, une supervision humaine est nécessaire lors de la conception de l'algorithme. D'autre part, une supervision humaine est nécessaire lors de l'utilisation en vie réelle de l'algorithme. Cette seconde supervision pourrait se matérialiser par la création de « *Collèges de Garantie Humaine* ». Ces structures réuniraient différents professionnels et représentants tels que « *médecins, professionnels de santé paramédicaux et représentants des usagers qui réviseraient des dossiers médicaux pour porter un regard humain sur les options thérapeutiques conseillées ou prises par l'algorithme* » (2).

### **1.2.2 Textes préparatoires au projet de loi relatif à la bioéthique**

Afin de bien visualiser les différents aspects de la garantie humaine, nous étudierons de façon chronologique son apparition progressive dans les textes préparatoires au projet de loi relatif à la bioéthique.

L'article 47 de la loi de bioéthique du 7 juillet 2011 prévoit un nouvel examen de l'ensemble de la loi par le Parlement (composé de l'Assemblée nationale et du Sénat) dans un

délai maximal de sept ans (14). De plus, cette révision doit désormais être précédée d'un débat public sous forme d'États généraux de la bioéthique. Cette nouvelle mission est confiée au CCNE (14). La dernière révision de la loi de bioéthique date de 2011. L'initiative de cette révision vient du Premier ministre et c'est pour cela qu'on parle de « projet de loi ». *A contrario*, si la demande émanait d'un sénateur ou d'un député, on parlerait de « proposition de loi ».

Une douzaine de textes au total ont été nécessaires pour préparer la révision de cette loi. Cinq d'entre eux abordent explicitement le principe de garantie humaine, à l'exception de la délibération n°2019-097 de la CNIL qui reprend son contenu sans nommer le principe :

- Les avis 129 et 130 du CCNE (5)
- Le rapport d'information déposé par la Mission d'information sur la révision de la loi relative à la bioéthique, enregistré à la Présidence de l'Assemblée nationale le 15 janvier 2019 (38)
- La délibération n°2019-097 du 11 juillet 2019 portant avis sur un projet de loi relatif à la bioéthique de la CNIL (45)
- L'étude d'impact du projet de loi relatif à la bioéthique de l'Assemblée nationale du 23 juillet 2019 (46)

#### 1.2.2.1 Les avis 129 et 130 du CCNE

Le principe d'une « garantie humaine » a été proposé par le groupe de travail à l'origine du rapport « *numérique et santé : quels enjeux éthiques pour quelles régulations* » commandé par le CCNE et rendu public en novembre 2018 ; le terme a été repris dans l'avis 129 du CCNE. Chapitre « numérique et santé ». Dans l'avis 129, le CCNE propose d'inscrire au niveau législatif le principe fondamental d'une garantie humaine du numérique en santé. Il définit ce principe de la façon suivante : « ***la garantie d'une supervision humaine de toute utilisation du numérique en santé, et l'obligation d'instaurer pour toute personne le souhaitant et à tout moment, la possibilité d'un contact humain en mesure de lui transmettre l'ensemble des informations la concernant dans le cadre de son parcours de soins*** » (5). La garantie humaine s'articule donc autour de deux paramètres : l'obligation d'information au patient et la supervision humaine.

Dans l'avis 130, rendu public le 29 mai 2019, le CCNE revient longuement sur l'importance de cette garantie humaine, qui implique que chaque patient soit informé de tout



recours à un traitement algorithmique de données massives. Le Comité estime « *qu'une garantie humaine des différentes étapes du processus de l'analyse des données est fondamentale.* » (47). Ces étapes sont les suivantes : la qualité des données, l'adéquation des traitements algorithmiques à la question posée, et la vérification sur un jeu de données indépendantes de la robustesse et de l'exactitude du résultat donné par l'algorithme. Pour le CCNE, cette garantie est d'autant plus importante que le préjugé d'infaillibilité et d'objectivité que l'on accorde généralement aux analyses fondées sur des modèles informatiques et mathématiques induit une confiance excessive qui peut accroître le risque d'erreur.

#### 1.2.2.2 Le rapport d'information par la Mission d'information sur la révision de la loi relative à la bioéthique

Ce rapport a été enregistré à la Présidence de l'Assemblée nationale le 15 janvier 2019. Auditionné le 23 octobre 2018, Monsieur David Gruson est favorable à l'introduction du principe de garantie humaine dans la loi. Il conviendra par la suite d'encourager « *l'émergence de techniques opérationnelles de recommandation de bonnes pratiques pour essayer de mettre concrètement en œuvre cette garantie humaine de l'IA.* » (38). Selon lui, ce rôle pourrait revenir à la Haute Autorité de santé (HAS).

Comment le principe de garantie humaine pourrait-il se matérialiser ? Trois propositions ont été formulées dans ce rapport (38) :

1) Par des procédés de vérification régulière des algorithmes : Monsieur David Gruson propose la création de « **collèges de garantie humaine** ». Ces collèges auraient pour mission d'étudier, à échéance régulière de 2-3 mois environ, et à l'échelle d'un établissement de santé ou d'un territoire plus large, plusieurs dizaines de dossiers médicaux où il aura été recouru à un algorithme d'aide au diagnostic, afin de vérifier la pertinence du diagnostic fourni par l'algorithme.

2) Par le recours à une télémedecine de garantie humaine pour exercer une supervision humaine en amont du traitement. La notion de **télémedecine** a été développée par la Société française de télémedecine (48).

3) Par la garantie de la **maîtrise de la décision finale** par le médecin. Le Conseil d'État suggère d'utiliser l'ancrage légal de la répression de l'exercice illégal de la médecine posé

par l'article L4161-1 du Code de la santé publique pour justifier l'interdiction d'un diagnostic établi uniquement par un système d'IA sans intervention conjointe d'un médecin (49). D'ailleurs, l'article 10 de la loi du 6 janvier 1978 dispose « *qu'aucune décision produisant des effets juridiques à l'égard d'une personne ne peut être prise sur le seul fondement d'un traitement automatisé des données.* » (50). À ce sujet, le Conseil d'État insiste sur la nécessité de rappeler de façon plus nette cette interdiction (10).

### 1.2.2.3 La délibération portant avis sur un projet de loi relatif à la bioéthique de la CNIL du 11 juillet 2019

La Commission a été saisie le 14 juin 2019 par le ministère des Solidarités et de la Santé, d'une demande d'avis portant sur quatre mesures du projet de loi relatif à la bioéthique. L'une de ces mesures vise à « *sécuriser la bonne information du patient lorsqu'un traitement algorithmique de données massives est utilisé à l'occasion d'un acte de soins* » (45). Cette mesure sera reprise à l'article 11 du projet de loi de bioéthique. Cet article insère un nouvel article L.4001-3 dans le CSP relatif aux garanties prévues en cas de recours à un traitement algorithmique de données massives par un professionnel de santé.

Quatre remarques sont formulées par la CNIL sur la rédaction de cet article 11 (45) :

- **L'ambiguïté du concept de « *traitement algorithmique de données massives* »** (TADM) utilisé pour statuer sur la situation individuelle d'un patient. Il conviendrait de distinguer le TADM en amont de la prise en charge individuelle, pour identifier des règles qui seront ensuite appliquées lors de la prise en charge d'un patient, dont les données seront soumises au système.

- **L'ambiguïté de l'obligation « *d'information des personnes concernées* »** : à quel moment le médecin doit-il informer le patient du recours à un TADM dans sa prise en charge ? Préalablement à l'utilisation du traitement ou lors de la délivrance des résultats ? La CNIL soutient la première option pour préserver la portée des articles L1111-2 et L1111-4 du CSP (51)(52). À la lecture de ces articles, le patient a le droit d'être informé sur son état de santé et des traitements qui lui sont proposés, et prend, avec le professionnel de santé et compte tenu des informations et des préconisations fournies, les décisions concernant sa santé.

- **L'ambiguïté du concept de « *communication des résultats* »** : La CNIL recommande que le projet de loi bioéthique distingue clairement le résultat brut issu de l'algorithme de l'appréciation, de la décision du professionnel de santé dans le cadre de la prise en charge du patient. Trois raisons appuient cette demande d'éclaircissement. D'abord, pour garantir l'accès de la personne concernée à l'ensemble des informations qui la concerne. Ensuite, pour que les principes d'indépendance et de liberté de prescription du professionnel de santé soient garantis. Enfin, pour garantir que la décision médicale ne se fonde pas exclusivement sur un traitement automatisé de données conformément aux dispositions des articles 22 du RGPD et 47 de la loi informatique et libertés (26)(50).

- **L'ambiguïté autour du « *paramétrage du traitement algorithmique* »**. La CNIL s'interroge sur :

1. Le rôle précis du professionnel de santé sur un tel dispositif, la nature des modifications qu'il pourra lui apporter (choix des données massives, options de traitement des données du patient, etc.). L'intervention humaine sur le paramétrage du traitement algorithmique ne pourrait-elle pas plutôt porter, non sur chaque décision individuelle, mais sur des séries de décisions ?

2. Les conséquences qui pourraient en résulter aussi bien pour le professionnel de santé (mise en jeu de sa responsabilité) que pour le patient (en cas d'erreur médicale ou de dysfonctionnement de l'outil par exemple).

Par conséquent, la CNIL propose d'une part, que le professionnel de santé renseigne lui-même dans l'outil, les données individuelles du patient dont il assure la prise en charge. D'autre part, qu'il soit permis au professionnel de santé de modifier le traitement algorithmique lui-même. La CNIL reconnaît qu'une telle pratique devrait faire l'objet d'un encadrement plus précis, car tout professionnel de santé ne dispose pas des compétences techniques nécessaires à de telles interventions.

1.2.2.4 L'étude d'impact du projet de loi relatif à la bioéthique de l'Assemblée nationale du 23 juillet 2019

Dans cette étude, l'Assemblée nationale semble répondre à deux questions fondamentales.

- Quelle est la place respective de l'algorithme, du patient et du professionnel de santé ?

L'Assemblée nationale alerte sur le risque que le caractère « scientifique/mathématique » de la conclusion produite par un algorithme paraisse devoir s'imposer de façon indiscutable aussi bien au professionnel de santé qu'au patient. De plus, « *la validité des conclusions produites par ces dispositifs peut toutefois être limitée par des défauts ou biais (volontaires ou non)* » (46).

➤ Comment assurer la maîtrise de la décision finale par le professionnel de santé ?

Le professionnel de santé doit être susceptible de remettre en perspective les résultats rendus par un dispositif d'IA par rapport aux autres informations, notamment cliniques, dont il peut disposer. Cela permet de « *garantir que la décision finale, qui s'appuie sur ces dispositifs, est prise par le professionnel de santé et le patient lui-même, notamment aux fins de respect du principe législatif de consentement éclairé aux soins* » (46).

#### 1.2.2.5 Les cinq objectifs de la garantie humaine

L'étude chronologique des textes préparatoires a permis de résumer le principe de garantie humaine autour de cinq objectifs poursuivis (46) :

- 1) **Rappeler qu'une décision médicale ne peut résulter du seul résultat d'un traitement automatisé.**
- 2) **Assurer la protection aussi bien du patient que du professionnel de santé.**
- 3) **Préciser l'utilisation de traitement algorithmique.**
- 4) **Conforter le droit à l'information du patient.**
- 5) **Garantir la traçabilité des données de santé utilisées, ainsi que l'accès par les professionnels de santé aux informations contenues dans ces dispositifs.**

#### 1.2.3 Article 11 du projet de loi relatif à la bioéthique

Le principe de garantie humaine de l'IA et du numérique en santé, introduit et porté par Ethik-IA depuis 2017, avait d'ores et déjà été reconnu dans les avis 129 et 130 du CCNE et dans l'article 11 du projet de loi de bioéthique. L'article 11 finit actuellement son parcours législatif devant le Parlement pour une entrée en vigueur en France à la fin de l'année 2021.

Aux termes de l'article 45, alinéa 1, de la Constitution française : « *Tout projet ou proposition de loi est examiné successivement dans les deux assemblées du Parlement en vue de l'adoption d'un texte identique.* » (53).

Le projet de loi relatif à la bioéthique se structure autour de sept axes :

**Titre I<sup>er</sup>** : élargit l'accès aux technologies déjà disponibles en matière de procréation tout en continuant de les encadrer.

**Titre II** : encourager la solidarité entre les personnes, notamment par le don d'éléments et produits du corps humain, sans renoncer au respect et à la protection des droits individuels.

**Titre III** : diffuser de nouveaux progrès scientifiques et technologiques dans le respect des principes éthiques.

**Titre IV** : soutenir une recherche libre et responsable, au service de la santé humaine.

**Titre V** : poursuivre l'amélioration de la qualité et de la sécurité des pratiques concernées par le champ bioéthique.

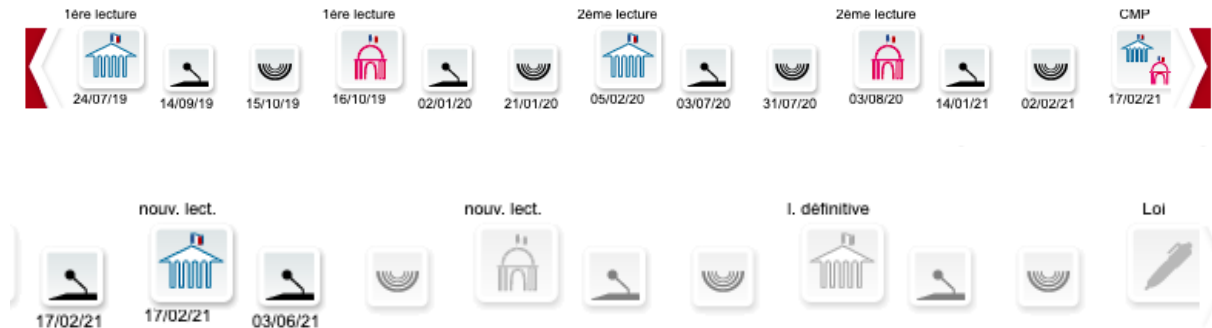
**Titre VI** : noter les mesures qui permettront d'installer une gouvernance bioéthique adaptée au rythme des avancées des sciences et des techniques.

**Titre VII** : prévoir plusieurs ordonnances.

**Le principe de garantie humaine du numérique en santé correspond à l'article 11 du Titre III du projet de loi relatif à la bioéthique.** En pratique, il apparaîtra sous la forme d'un nouvel article du Code de santé publique une fois la loi promulguée, à savoir **l'article L.4001-3 CSP.**

Figure 5 : Les étapes de la discussion du projet de loi relatif à la bioéthique

## Les étapes de la discussion :



Source : [www.sénat.fr](http://www.sénat.fr), 07/06/2021, <https://www.senat.fr/dossier-legislatif/pjl19-063.html>

(Autorisation diffusion en annexe)

À ce stade de la rédaction du mémoire, quatre versions de l'article 11 du projet de loi relatif à la bioéthique ont été présentées. Elles correspondent aux différentes étapes de la navette parlementaire : deux lectures du texte par chaque chambre.

Le travail qui va suivre nous a permis d'analyser l'évolution du principe de garantie humaine à la lumière des différentes versions du projet de loi relatif à la bioéthique.

Il est important de préciser que les prochains paragraphes relèvent de notre interprétation grâce à une analyse linéaire des quatre versions de l'article 11 du projet de loi relatif à la bioéthique. Le surlignage jaune a été ajouté par nos soins pour mettre en avant les enjeux et modifications qui nous semblent importants.

Le 15 octobre 2019, l'Assemblée nationale a adopté, en première lecture, l'article 11 du projet de loi relatif à la bioéthique. Cet article introduit un nouvel article L.4001-3 dans le Code de la santé publique.

L'article 11 a été modifié à quatre reprises suivant le schéma de la navette parlementaire. Nous allons à présent en étudier les différents changements apportés à chaque relecture.

- 1<sup>re</sup> version (Assemblée nationale)

L'article L.4001-3 du CSP, dans sa version initiale, se compose de trois alinéas.

- Alinéa 1 : « *Lorsque, pour des actes à visée préventive, diagnostique ou thérapeutique, est utilisé un traitement algorithmique de données massives, le professionnel de santé qui communique les résultats de ces actes informe la personne de cette utilisation et des modalités d'action de ce traitement.* »

Dans son premier alinéa, l'article L.4001-3 du CSP aborde la question de l'utilisation d'un traitement algorithmique de données massives pour des actes à visée préventive, diagnostic ou thérapeutique. Ce TADM se réfère à l'utilisation d'une intelligence artificielle. En effet, tout traitement algorithmique n'emploie pas systématiquement de l'IA. Autrement dit, cet article encadre l'emploi d'outils d'IA dans le dépistage de la maladie (« actes à visée préventive »), dans l'identification de la maladie (« actes à visée diagnostique ») et dans le traitement de la maladie (« actes à visée thérapeutique »).

Au stade de la communication des résultats médicaux, le professionnel de santé doit informer le patient que pour obtenir ces résultats médicaux, il a eu recours à un outil utilisant de l'IA. L'information à la personne concernée intervient seulement au stade de la délivrance des résultats, et non préalablement à l'utilisation de l'outil.

- Alinéa 2 : « *L'adaptation des paramètres d'un traitement mentionné au I pour des actions à visée préventive, diagnostique ou thérapeutique concernant une personne ne peut être réalisée sans l'intervention d'un professionnel de santé.* »

Dans son deuxième alinéa, l'article L.4001-3 du CSP aborde le principe de garantie humaine sans le mentionner explicitement. Au stade de l'utilisation de l'outil d'IA, le professionnel de santé doit lui aussi intervenir dans l'adaptation des paramètres de cet outil aux côtés d'autres acteurs tels que certainement les techniciens ou informaticiens. Au stade de la conception de l'outil d'IA, le texte n'aborde pas la question d'une possible intervention du professionnel de santé au sein d'une équipe très hétérogène qui pourrait se composer, en plus des concepteurs de l'outil, de spécialistes de la pathologie pour laquelle l'outil serait utilisé et de spécialistes de l'éthique médicale.

- Alinéa 3 : « La **traçabilité** des actions d'un traitement mentionné au I et des données ayant été utilisées par celui-ci est assurée et les informations qui en résultent sont accessibles aux professionnels de santé concernés. »

Ce dernier alinéa complète le premier alinéa de l'article L.4001-3 du CSP. En effet, le professionnel de santé, pour être capable de comprendre et ensuite expliquer au patient les résultats proposés par l'outil, doit être en mesure d'accéder à l'historique des actions menées et des données utilisées par l'outil. Cet alinéa s'applique uniquement aux professionnels de santé concernés par les résultats de la machine.

- 2<sup>e</sup> version (Sénat)

Le 4 février 2020, le Sénat en première lecture, a apporté des modifications à l'article 11 du projet de loi relatif à la bioéthique. L'article L.4001-3 du CSP se compose désormais de cinq alinéas.

- Alinéa 1 : « Lorsque, pour des actes à visée préventive, diagnostique ou thérapeutique, le professionnel de santé envisage de recourir à un traitement algorithmique, il en **informe préalablement** le patient et lui explique sous une **forme intelligible** la manière dont ce traitement serait mis en œuvre à son égard. Seules **l'urgence** et **l'impossibilité** d'informer peuvent y faire obstacle. »

Cette version ne mentionne pas expressément l'emploi de l'intelligence artificielle. Elle se limite au recours à un traitement algorithmique sans préciser si dans celui-ci, est incorporée de l'IA ou non.

Contrairement à la première version, le professionnel de santé doit informer le patient de son intention de recourir à un traitement algorithmique, pour des actes à visée préventive, diagnostic ou thérapeutique, en amont de son utilisation. Autrement dit, le professionnel de santé ne peut recourir à ce traitement algorithmique sans en avoir informé avant le patient.

Par l'emploi du terme « envisage », le professionnel de santé se situe au stade de la réflexion, c'est-à-dire qu'il va être en mesure de choisir d'utiliser ou non un traitement algorithmique. Dans la version antérieure, la question de la liberté de choix du professionnel



de santé n'était même pas envisagée puisqu'on se situait au stade où ce dernier avait déjà utilisé l'algorithme.

Cette version précise la manière dont l'information sur le fonctionnement de l'algorithme doit être délivrée par le professionnel de santé : elle doit être compréhensible par le patient.

Le Sénat apporte deux exceptions à cette obligation d'information : l'urgence et l'impossibilité d'informer. Dans ces cas-là, le professionnel de santé pourrait informer le patient de son utilisation à un traitement algorithmique postérieurement à celle-ci comme dans la première version de l'article 11.

- Alinéa 2 : « La **saisie d'informations** relatives au patient dans le traitement algorithmique se fait sous le **contrôle** du **professionnel de santé** qui a recourt audit traitement ».

L'ancien alinéa 2 de la première version est remplacé par un nouvel alinéa. Le professionnel de santé supervise les informations concernant son patient qui sont saisies dans l'algorithme.

- Alinéa 3 : « **Aucune décision médicale** ne peut être prise sur le **seul fondement** d'un traitement **algorithmique** ».

L'ancien alinéa 3 est remplacé par un nouvel alinéa. Cet alinéa ne précise pas explicitement sur qui repose la maîtrise de la décision finale, si ce n'est que les résultats de la machine ne peuvent pas passer directement de la machine au patient.

- Alinéa 4 : « Les **concepteurs** d'un traitement algorithmique mentionné au premier alinéa s'assurent de la transparence du fonctionnement de l'outil pour ses utilisateurs. »

Cet alinéa est nouveau par rapport à la version précédente. La transparence impose la traçabilité des opérations, mais aussi potentiellement des démarches visant à assurer l'explicabilité des algorithmes.

- Alinéa 5 : « Un décret en Conseil d'État, pris après avis motivé et publié de la Commission nationale de l'informatique et des libertés, précise les modalités d'application du présent article, notamment la mise en œuvre de l'information du patient,

*les conditions d'utilisation du traitement algorithmique par les professionnels de santé et celles dans lesquelles la transparence du fonctionnement dudit traitement est assurée par son concepteur. »*

Cet alinéa est nouveau par rapport à la version précédente. Un décret en Conseil d'État est désormais prévu pour la mise en œuvre de l'article 11.

- 3<sup>e</sup> version (Assemblée nationale)

Le 31 juillet 2020, l'Assemblée nationale en deuxième lecture a apporté des modifications à l'article 11 du projet de loi relatif à la bioéthique. L'article L.4001-3 du CSP est réduit à trois alinéas.

- Alinéa 1 : *« Lorsque, pour des actes à visée préventive, diagnostique ou thérapeutique, est utilisé un traitement algorithmique dont l'apprentissage est réalisé à partir de données massives, le professionnel de santé qui décide de cette utilisation s'assure que la personne concernée en a été informée au préalable et qu'elle est, le cas échéant, avertie de l'interprétation qui en résulte. »*

Par rapport à la version précédente de l'article (1<sup>re</sup> lecture, Sénat), l'Assemblée nationale rappelle de nouveau que l'information du patient concerne le recours à l'IA. Le choix du professionnel de santé de recourir à cette utilisation est clairement affirmé. Par contre, la façon dont cette information doit être délivrée au patient n'est plus précisée. De même, les deux exceptions à cette information (« l'urgence » et « l'impossibilité ») ont été supprimées.

Si cette information au patient sur le recours à l'IA est de nouveau prévue en amont de son utilisation par le professionnel de santé ; il est toutefois envisagé qu'elle puisse être délivrée au moment des résultats.

- Alinéa 2 : *« La traçabilité des actions d'un traitement mentionné au I et des données ayant été utilisées par celui-ci est assurée et les informations qui en résultent sont accessibles aux professionnels de santé concernés. Les données sont traitées et partagées dans les conditions prévues à l'article L. 1111-8. »*

L'Assemblée nationale a repris la formulation qu'elle avait proposée en première lecture, à l'alinéa 3. Elle ajoute un article du Code de la santé publique auquel se référer pour le traitement et le partage des données de santé.

Alinéa 3 : « Un **arrêté du ministre chargé de la santé** établit, après **avis de la Haute Autorité de santé**, la liste des types de traitements algorithmiques qui font l'objet de l'information mentionnée au I. Il détermine, après **avis de la Commission nationale de l'informatique et des libertés**, pour chaque type de traitements, la nature et la durée de conservation des actions et des données dont la traçabilité est prévue au II. »

Cet alinéa est nouveau par rapport aux versions précédentes. Désormais, la mise en œuvre de l'article 11 est prévue, non plus par un décret du Conseil d'État, mais par un arrêté du ministre chargé de la santé après un avis de HAS et un avis de la CNIL.

- 4<sup>e</sup> version (Sénat)

Le 3 février 2021, le Sénat en deuxième lecture a apporté des modifications à l'article 11 du projet de loi relatif à la bioéthique. L'article L.4001-3 du CSP se compose désormais en quatre alinéas.

- Alinéa 1 : « *Lorsque, pour des actes à visée préventive, diagnostique ou thérapeutique, est utilisé un traitement algorithmique dont l'apprentissage est réalisé à partir de données massives, le professionnel de santé qui décide de cette utilisation s'assure que la personne concernée en a été informée au préalable et qu'elle est, le cas échéant, avertie de l'interprétation qui en résulte.* »

Cet alinéa est identique à la version précédente.

- Alinéa 1 bis : « *Aucune décision médicale ne peut être prise sur le seul fondement d'un traitement **mentionné au I.*** »

Cet alinéa est une reprise de l'alinéa 3 de la deuxième version de l'article proposée en première lecture par le Sénat : « *Aucune décision médicale ne peut être prise sur le seul fondement d'un traitement* ». Toutefois, dans cette nouvelle version, le Sénat opère un renvoi au premier alinéa pour encadrer la décision médicale ; ce qui n'était pas le cas en première lecture.

- Alinéa 2 : « Un **principe de garantie humaine** s'applique à ces traitements algorithmiques. La mise en œuvre de ce principe est notamment assurée par le **fabricant** dans les conditions prévues dans le cadre de la mise sur le marché du traitement algorithmique. »

Pour la première fois, l'expression de « *principe de garantie humaine* » est clairement employée. Contrairement aux versions précédentes, la garantie humaine n'est plus abordée sous l'angle du professionnel de santé ou du concepteur, mais sous l'angle du fabricant.

Cet alinéa reprend l'amendement de Monsieur Alain Milon adopté en 2<sup>de</sup> lecture au Sénat après avis favorable du Gouvernement (54).

- Alinéa 3 : « *Un arrêté du ministre chargé de la santé établit, après avis de la Haute Autorité de santé et de la Commission nationale de l'informatique et des libertés, la liste des types de traitements algorithmiques qui font l'objet de l'information mentionnée au I. Cette liste est régulièrement mise à jour.* »

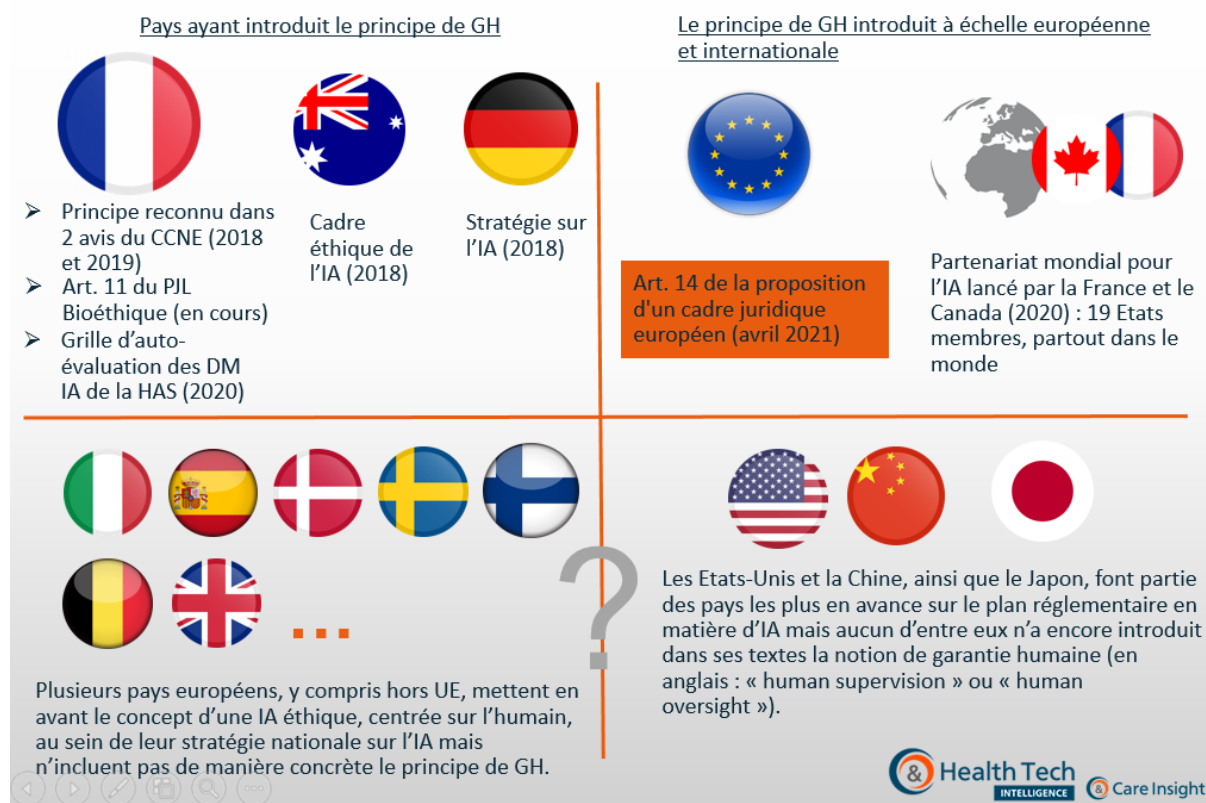
La rédaction retenue pour la mise en œuvre de l'article 11 a été simplifiée.

## 1.2.4 Introduction du principe de garantie humaine à l'international

Le cadre juridique de la garantie humaine a été repris par de nombreuses instances.

Figure 6 : Introduction du principe de garantie humaine à l'international

### IA : introduction du principe de Garantie Humaine à l'international



Source : Health & Tech Intelligence - Care Insight, 01/06/2021 (autorisation diffusion en annexe)

Au cours de l'année 2020-2021, le principe de garantie humaine a reçu, en plus de la révision bioéthique, des concrétisations dans quatre autres cadres significatifs (42):

- La garantie humaine a été intégrée à la grille d'auto-évaluation des dispositifs médicaux intégrant de l'IA préalablement à leur admission au remboursement publié en octobre 2020 par l'HAS (55) ;
- Le principe a été repris dans le Livre Blanc sur l'IA publié par la Commission européenne le 19 février 2020(56) ;
- Le principe est introduit à l'article 14 du projet de règlement sur l'intelligence artificielle diffusé le 21 avril 2021 par la Commission européenne.(7) L'article 14 est plus détaillé que

l'article 11 du projet de loi bioéthique. Il reprendra les deux polarités : information de l'utilisateur sur l'utilisation du recours à l'IA et la supervision humaine. La Commission européenne a ajouté un élément nouveau, qui ne figurait pas dans la proposition initiale d'Ethik-IA : une capacité d'interruption de l'IA à tout moment (41) ;

- Le principe de garantie humaine est repris dans le cadre task-force dédiée par l'OMS à la régulation de l'IA en santé en vue de l'émission d'une recommandation générale en juillet 2021.

En somme, le modèle français de la bioéthique est construit autour de trois principes : la dignité, la liberté et la solidarité. Le développement exponentiel des données assorti de l'essor de l'IA en santé nous invite inéluctablement à repenser notre modèle ; au risque de se voir imposer des modèles étrangers en contradiction avec nos valeurs éthiques (57). « *Les perspectives ouvertes par l'essor de l'intelligence artificielle et le développement exponentiel de traitements de données de santé comme par le phénomène des mégadonnées suscitent des questions nouvelles, porteuses d'espoirs médicaux sans précédent, mais également de menaces pour la protection des personnes et de leur vie privée et même pour l'évolution de l'espèce humaine* » (17).

« *S'ouvrir résolument à l'innovation et essayer d'en réguler les enjeux éthiques au fil de son application : c'est le sens du principe de garantie humaine de l'IA en santé* » (2). Venu de la fiction, ce principe a été repris dans les travaux de préparation du groupe de travail du CCNE co-dirigé par Monsieur David Gruson pour la révision de la loi de bioéthique. L'inscription du principe de garantie humaine à l'article 11 du projet de loi de bioéthique et à l'article 14 du projet de règlement de l'IA représente de nouvelles étapes fortes de reconnaissance et de concrétisation de ce principe général de régulation positive de l'IA en santé.

## 2 : PROBLÉMATIQUE ET HYPOTHÈSES

---

### 2.1 PROBLÉMATIQUES

Si l'étude du sujet peut paraître large, celle-ci est doublement limitée. D'une part, le sujet se cantonnera aux applications de l'IA dans le domaine de la santé. D'autre part, on s'interrogera principalement sur la période de 2017 à nos jours, puisque le principe de garantie humaine du numérique en santé n'aurait jamais été proposé avant en France, et ne sera applicable que lorsque la nouvelle loi de bioéthique française sera promulguée.

Historiquement, la relation médecin-malade était un contrat qui s'établissait entre deux personnes : le malade qui sollicitait les meilleurs soins et le médecin investi d'une mission dans laquelle la confiance devait être établie (58). Désormais, un troisième acteur, l'IA, va progressivement intégrer cette relation grâce au développement des outils algorithmiques.

L'article 10 de la loi du 6 janvier 1978, actuellement en vigueur, pose clairement ce principe : « aucune décision produisant des effets juridiques à l'égard d'une personne ne peut être prise sur le seul fondement d'un traitement automatisé des données » (50). Or cette loi ne consacre pas explicitement l'interdiction d'un diagnostic établi uniquement par un système d'IA sans intervention d'un médecin (10).

Les thèmes « Intelligence artificielle et robotisation » étaient absents des dernières révisions des lois de bioéthiques. En 2018, ils intègrent pour la première fois, le périmètre des États généraux de la bioéthique (1).

Il faudra comprendre comment l'IA interfère dans le colloque singulier, puisque le sujet suppose implicitement que l'IA va progressivement exercer un rôle précis dans la détermination du traitement thérapeutique proposé au patient grâce à l'utilisation d'outils d'aide au diagnostic ou d'aide à la décision. La question de recherche pose alors le cadre suivant : il ne s'agit pas de savoir si l'IA joue un rôle dans le domaine médical, mais plutôt d'identifier quels sont ce rôle et ce lien qui unissent l'IA au principe de garantie humaine du numérique en santé.

De plus, le projet de révision de la loi de bioéthique, dans son article 11, intègre le principe de garantie humaine du numérique en santé, mais son contenu ne cesse d'évoluer au fur et à mesure des débats parlementaires. Si le législateur donne sans le moindre doute, une valeur symbolique au principe de garantie humaine en l'intégrant dans la nouvelle loi de bioéthique, ce principe aura-t-il pour autant une réelle portée législative ?

Ainsi, le sujet met en lumière un paradoxe : en théorie, aucune décision produisant des effets juridiques à l'égard d'une personne ne peut être prise sur le seul fondement d'un traitement automatisé de données. D'où vient ce besoin de consacrer au niveau législatif un principe de garantie humaine du numérique en santé ? Quelles conséquences cette consécration aura-t-elle sur le développement des applications de l'IA en santé ? Comment travailler sur un cadre juridique en cours de création qui a vocation à s'appliquer sur des outils algorithmiques encore peu utilisés aujourd'hui ? Comment concilier la demande croissante de régulation juridique émanant de la Société avec l'évolution rapide des innovations scientifiques et technologiques ? (9)

## 2.2 QUESTION DE RECHERCHE ET OBJECTIFS

Notre axe de recherche s'articulera autour de la problématique suivante : **comment le principe de « garantie humaine » du numérique en santé peut-il s'inscrire dans le développement des applications de l'intelligence artificielle ?**

L'objectif principal est de mieux comprendre le principe de garantie humaine et ses interactions avec l'IA en santé.

L'objectif secondaire est de recueillir une vision des concepteurs d'outils d'IA sur ce principe : leurs représentations sur les questionnements des législateurs autour des enjeux d'IA et des données de santé, ainsi que leur avis sur leur potentielle responsabilité engagée en tant que concepteurs d'outils d'IA.

## 2.3 HYPOTHÈSES DE RECHERCHE

La première hypothèse était celle d'une notion peu connue du principe de garantie humaine du numérique en santé, due à son caractère nouveau et à la complexité des évolutions législatives.



La seconde hypothèse avait trait à la responsabilité : les concepteurs d'outils d'aide à la décision ne se considèrent actuellement pas responsables en cas d'erreur médicale et la faute professionnelle est imputée aux soignants.

## 3 : MÉTHODOLOGIE

---

« *Il y a quelque chose de mystérieux dans la rencontre d'une sensibilité (celle du chercheur) et d'une expérience (celle du participant à la recherche) et cela doit être honoré et respecté.* » - Paillé et Mucchielli (59)

### 3.1 CONSTITUTION D'UNE BIBLIOGRAPHIE

La constitution d'une bibliographie fut une étape primordiale de l'étude. Dans un domaine aussi vaste et diversifié qu'est celui de l'éthique médicale et de la recherche, il nous paraissait intéressant de pouvoir s'appuyer essentiellement sur une étude chronologique des textes préparatoires à la loi de bioéthique et sur les différentes versions de l'article 11 du projet de loi de bioéthique, pour comprendre la genèse et l'évolution du principe de garantie humaine jusqu'à sa consécration légale.

Cette bibliographie a été complétée par de nombreuses recherches et diverses recommandations. L'objectif de ces recherches complémentaires était d'acquérir autant que possible, une certaine variété des sources, pour ne pas se contenter de sources exclusivement juridiques.

Zotero, PubMed, Dalloz, Lexis360, HAL, *Ex Libris Discovery* sont autant de sites qui ont permis l'identification des mots clefs pertinents, la recherche d'ouvrages, de revues, d'articles en rapport avec la recherche.

Il convient toutefois de reconnaître que la bibliographie est composée majoritairement de textes de nature juridique.

Bien sûr, la littérature ne pouvait se suffire à elle-même. L'objectif de notre projet n'était pas de faire une revue de la littérature. Aussi, la mise en place de plusieurs entretiens semi-dirigés avec des concepteurs d'outils d'aide à la décision puis avec des professionnels de santé familiers de l'IA en santé semblait compléter de façon opportune la théorie et la pratique.

## **3.2 ÉTUDE QUALITATIVE RÉALISÉE**

Il s'agissait d'une étude qualitative observationnelle et monocentrique.

Notre étude s'intéressait d'abord aux recueils des avis des personnes concernées par le principe de garantie humaine (concepteurs d'outils d'IA, professionnels de santé, patients, juristes, etc.). Nous souhaitions explorer, à travers leurs regards, les tensions, les complémentarités et les écarts avec les motivations sous-jacentes aux discussions autour de la définition d'une garantie humaine et de son inscription en Droit.

### **3.2.1 Échantillon composé de concepteurs d'outils d'IA et de professionnels de santé familiers de l'IA**

De ce fait, notre première intuition a été de nous concentrer sur les concepteurs d'outils d'IA, qui représentaient l'un des premiers maillons du développement de l'IA en santé. Par « concepteurs », on peut entendre des cliniciens ou chercheurs impliqués dans la démarche de développement d'outils d'IA.

Nous avons été associés au projet de définition d'un algorithme d'IA pour l'accident vasculaire cérébral (RHU BOOSTER). De fait, nous avons d'abord songé à assister à des réunions de ce projet BOOSTER pour évaluer la place du principe de garantie humaine et les représentations de ses enjeux, dans leurs discussions.

Cela n'ayant finalement pas été possible, en lien avec des obstacles d'organisations et de confidentialité d'un projet en cours. Nous avons fait le choix de mener des entretiens semi-directifs à plusieurs et individuels avec des concepteurs d'outils d'IA. L'objectif était de mieux comprendre le principe de garantie humaine, ses interactions avec l'IA en santé et de recueillir une vision des concepteurs de ces outils sur ce principe : leurs représentations sur les questionnements des législateurs autour des enjeux d'IA, ainsi que leur avis sur leur potentielle responsabilité engagée en tant que concepteurs d'outils d'IA en cas d'erreur médicale.

Progressivement, notre échantillon s'est élargi aux professionnels de santé familiers de l'IA en santé. Il nous semblait intéressant de pouvoir compléter, et par la suite confronter,

la compréhension et l'avis des concepteurs avec celui des professionnels de santé, sur le principe de garantie humaine.

### **3.2.2 Collecte d'informations matérialisées par des entretiens semi-directifs**

#### **3.2.2.1 Choix des entretiens semi-directifs**

Pour mémoire, un entretien peut être défini comme « *une méthode de collecte d'informations qui se situe dans une interaction entre un intervieweur et un interviewé en vue de partager un savoir expert et de dégager une compréhension d'un phénomène* » (60).

La méthodologie de l'entretien semi-directif consiste, pour cette étude, d'une part, de permettre aux concepteurs d'outils d'IA et professionnel de santé, d'exprimer les enjeux professionnels, juridiques et humains de l'IA au cœur de leur pratique. D'autre part, d'orienter la personne entretenue vers certains sujets, tout en la laissant par la suite libre de s'exprimer (61).

#### **3.2.2.2 Élaboration d'une grille d'entretien**

Les questions posées aux professionnels étaient divisées en deux catégories : d'abord, quatre questions larges sur l'application de l'IA en santé, puis quatre questions spécifiques autour du principe de garantie humaine.

Le dispositif d'entretien était le suivant :

- Quelle est votre expérience des outils d'aide à la décision utilisant de l'IA ?
- Quels avantages attendez-vous de ces outils ? Y voyez-vous des limites ?
- Quels sont les objectifs poursuivis par la création de ces outils ?
- Quelles questions posent ces outils ?
  
- Connaissez-vous le principe de garantie humaine du numérique en santé ?
- La garantie humaine vous paraît-elle répondre à vos interrogations sur les applications de l'IA en santé ?

- L'envisagez-vous comme : une source de limitation à la conception de ces outils ? Comme un cadre de protection de création et de production ?
- Ce principe de garantie humaine vous paraît-il pertinent ? Quel système idéal imagineriez-vous ? Comment en pratique cela pourrait-il prendre forme ?

La grille d'entretien complète est disponible en annexe 3.

### 3.2.2.3 Déroulement des entretiens

L'idée initiale était de réaliser ces entretiens en présentiel, mais en raison de l'épidémie actuelle, la plupart ont été réalisés en distanciel, via une application de visioconférence (Zoom).

Au total, sept entretiens ont été menés : six entretiens avec un participant par entretien, pour une durée d'une trentaine de minutes en moyenne ; et un entretien avec trois participants d'une heure et demie.

Les participants ont été contactés par mail et ont tous reçu en pièce jointe une lettre d'information concernant le protocole de recherche. Le mail type de prise de contact et la lettre d'information sont disponibles aux annexes 4 et 5.

Si l'on s'intéresse un peu plus aux détails, les entretiens comportaient :

- 1) Une annonce, c'est-à-dire la présentation de l'interviewer, celle de sa recherche et l'entrée en matière.
- 2) Les questions issues de la grille d'entretien sus évoquées.
- 3) Si la personne interrogée ne connaissait pas le principe de garantie humaine, une brève définition de notre compréhension de ce principe lui était donnée, afin de la mettre en contexte. Il était important pour nous de souligner à ce stade de l'entretien que l'objectif était de recueillir la vision de la personne interrogée sur ce qu'elle comprenait de ce principe grâce à ses connaissances de l'IA en santé.
- 4) Entre ces questions, des relances étaient faites lorsque celles-ci s'avéraient être nécessaires : il s'agissait alors d'entretenir la dynamique de l'entretien en reprenant parfois les propos des intervenants pour les encourager à développer davantage, à préciser certaines notions ou certaines sensations. Elles avaient alors pour effet « *d'inciter l'interlocuteur à faire un retour réflexif sur ce qu'il venait de dire : à*

*argumenter sa prise de position, à l'affirmer plus nettement ou à la nuancer, à se situer par rapport à ce qui lui apparaît comme une convergence ou une contradiction possible de ses propos » (5).*

- 5) En définitive, la conclusion passait par des remerciements et par l'évocation de nouvelles rencontres potentielles, éventuellement destinées à compléter ou préciser ultérieurement les propos recueillis.

L'enregistrement était lancé après les présentations d'usage et la déclaration d'anonymisation, puis coupé après les remerciements. Ils étaient par la suite intégralement retranscrits, tant dans le contenu littéral des propos qu'avec une mention visuelle des hésitations, des rires, des silences ou encore des exclamations.

#### 3.2.2.4 Élaboration d'une vignette clinique

Selon la HAS, *« l'appellation vignette clinique est utilisée pour désigner un cas clinique construit comme un scénario séquencé et accompagné d'une ou plusieurs questions qui explorent ce qu'un professionnel de santé ferait dans une situation semblable » (62).*

Pour notre recherche, nous avons fait le choix d'ajouter une vignette clinique. Le scénario était le suivant : dans un futur proche, un médecin du SAMU suivait les recommandations d'un outil d'IA conçu spécialement pour les accidents vasculaires cérébraux (AVC). Le patient décédait dans l'ambulance. La vignette est reproduite en annexe 3.

Grâce à l'aide de Frédéric Huynh, de formation médicale, et les observations formulées par d'autres camarades de classe, nous avons pu élaborer une vignette clinique validée par le Professeur Marie-France Mamzer. En effet, ma formation initiale en droit ne me permettait pas de la construire seule.

Cette vignette fut conçue pour les entretiens menés avec les participants du RHU BOOSTER. Elle était proposée en fin d'entretien et faisait partie de l'enregistrement. La vignette n'était pas proposée au reste de l'échantillonnage, car elle reprenait une situation médicale d'un AVC et il nous semblait donc plus pertinent d'interroger les participants du projet RHU BOOSTER plus familiers du sujet.

L'objectif de cette vignette était de mettre notre interlocuteur dans une situation que nous aurions construite au préalable, pour qu'il prenne une décision et nous fasse part de ses impressions. Nous voulions évaluer si les réponses aux questions précédentes étaient mobilisées à nouveau dans l'échange permis par la vignette, ou si l'interlocuteur soulevait de nouvelles tensions éthiques, professionnelles ou juridiques qui n'avaient pas été évoquées précédemment.

### **3.2.3 Aspects éthiques et réglementaires**

Trois aspects sont à préciser ici : le consentement éclairé des participants, la confidentialité et les autorisations réglementaires (63).

Les participants devaient avoir consenti librement à leur participation en toute connaissance de cause. Le plus simple et le plus sûr fut de joindre une lettre d'information dans le mail de recrutement, les informant de leurs droits (notamment de correction et de rétractation), et leur garantissant l'anonymat et la confidentialité.

Les propos recueillis lors des entretiens, les enregistrements audios se sont faits via l'enregistreur vocal de mon ordinateur personnel ou via le logiciel Zoom. Les données ont été extraites des enregistrements et ont été stockées sur mon ordinateur sécurisé par un mot de passe personnel. Pour éviter la perte des données, celles-ci ont été stockées sur un réseau sécurisé, dans la mesure du possible. Elles ont aussi été sauvegardées sur un disque dur dans un tiroir fermé à clef. Les données furent pseudonymisées et retranscrites par écrit. Les enregistrements, une fois retranscrits, ont été alors détruits. Les données recueillies sont des données à caractère personnel, mais pas des données de santé.

Concernant les durées de conservations des données utilisées, pour ce qui concerne les données issues des enregistrements audio, elles ont été conservées jusqu'à retranscription. Pour ce qui est des retranscriptions des entretiens, elles seront conservées jusqu'à publication.

En l'absence de traitement de données à caractère personnel, il ne fut pas nécessaire de procéder à des déclarations auprès de la Commission Nationale de l'Informatique et des Libertés (CNIL) et de l'Institut National des Données de Santé (INDS).

En effet, les données recueillies furent anonymisées conformément à la Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (2) et aux

recommandations de la CNIL. Grâce à un procédé de randomisation les noms, prénoms, noms d'établissements, genre des membres entretenus, furent supprimés.

Avant d'entamer les entretiens, notre recherche fut soumise à une demande d'avis au Comité des Recherches (CERAPHP Centre) qui valida notre recherche. De plus, nous avons envoyé une demande par mail au délégué à la protection des données (DPO) de l'Université de Paris pour déclarer nos données au Registre des traitements de données à caractère personnel mis en œuvre par l'Université de Paris.

### **3.2.4 Méthodes et déroulement des analyses de contenu**

#### **3.2.4.1 Choix d'analyse de contenu orienté vers la phénoménologie interprétative**

L'approche par analyse interprétative phénoménologique a été développée par Smith à partir des années 90 en Grande-Bretagne dans le domaine de la psychologie (64). L'objectif de cette approche est d'explorer la manière dont une expérience de vie a été ressentie et comprise par la personne qui l'a vécue. Le recrutement des participants était donc restreint à une population ayant vécu ou vivant le phénomène étudié. L'échantillonnage au sein de cette population était ciblé et homogène, c'est-à-dire que le chercheur devait s'astreindre à ne sélectionner que des personnes ayant en commun le vécu du même phénomène. Toutefois, il était recommandé de faire varier les autres caractéristiques de cette population (âge, genre, etc.). Le sens du travail de recherche se construit à partir du sens que chaque participant donna de son vécu (63).

Après avoir bien identifié la méthodologie, il convenait de l'adapter en pratique au moyen d'une grille d'analyse.

#### **3.2.4.2 Création de la grille d'analyse de contenu**

L'analyse était indépendante de chaque entretien et organisée en thèmes, regroupés en « thèmes superordonnés » et leur articulation. Les thèmes superordonnés correspondent à la recherche du sens commun à tous les participants par la synthèse des catégories conceptuelles issues de chaque entretien (63).



## 4 : RÉSULTATS

---

### 4.1 DÉROULEMENT DE L'ÉTUDE

L'étude s'est déroulée comme il était prévu, entre mars et avril 2021, au sein du RHU BOOSTER pour recueillir les visions des concepteurs d'outils algorithmiques utilisant de l'IA, sur le principe de garantie humaine. Par « concepteurs », nous entendons des cliniciens ou chercheurs impliqués dans la démarche de développement de ces algorithmes. L'étude a été complétée par plusieurs entretiens en dehors du RHU BOOSTER, toujours auprès de concepteurs, mais aussi de professionnels de santé familiers de l'IA. Ce choix se justifiait par la volonté de recueillir plusieurs visions. Les participants ont été contactés par mail, afin d'organiser l'entretien au mieux, dans le respect des conditions sanitaires et les contraintes organisationnelles, en lien avec l'épidémie du COVID-19. C'est pourquoi, si les premiers entretiens se sont déroulés en présentiel, les suivants étaient par visioconférence, à l'aide du logiciel Zoom. Onze personnes ont été contactées : neuf d'entre elles ont répondu de manière favorable et deux n'ont pas répondu. Au total, sept entretiens avec neuf personnes ont eu lieu dans les conditions prévues par le protocole de recherche.

La durée moyenne des entretiens était de trente minutes ; nous avons prévu quarante-cinq minutes en moyenne. L'entretien le plus court a été de vingt-cinq minutes, le plus long d'une heure et dix minutes.

Nous avons rencontré quelques difficultés au départ pour recruter des participants, car nous ne sommes ni du milieu médical ni du milieu informatique. Cette difficulté a été surmontée de trois façons. Premièrement, nos directeurs de mémoire nous ont aidés à établir les premiers contacts avec des concepteurs de leur réseau professionnel. Ensuite, nous avons pu bénéficier du réseau de notre Master 2, en contactant nous-mêmes certains intervenants des unités d'enseignements. Enfin, les premiers participants nous ont mis en contact avec des concepteurs et professionnels de santé.

Le premier entretien en groupe était très complexe par la technicité des échanges entre les participants. Suite à cet entretien, nous avons revu notre stratégie des entretiens et nous avons décidé de privilégier les entretiens individuels. De plus, en début d'entretien, notre formation initiale en Droit était soulignée pour justifier la demande d'effort de simplicité dans leurs réponses à venir. Les participants ont tous été compréhensifs et les entretiens se sont déroulés sans autre difficulté.

## **4.2 POPULATION DES CONCEPTEURS**

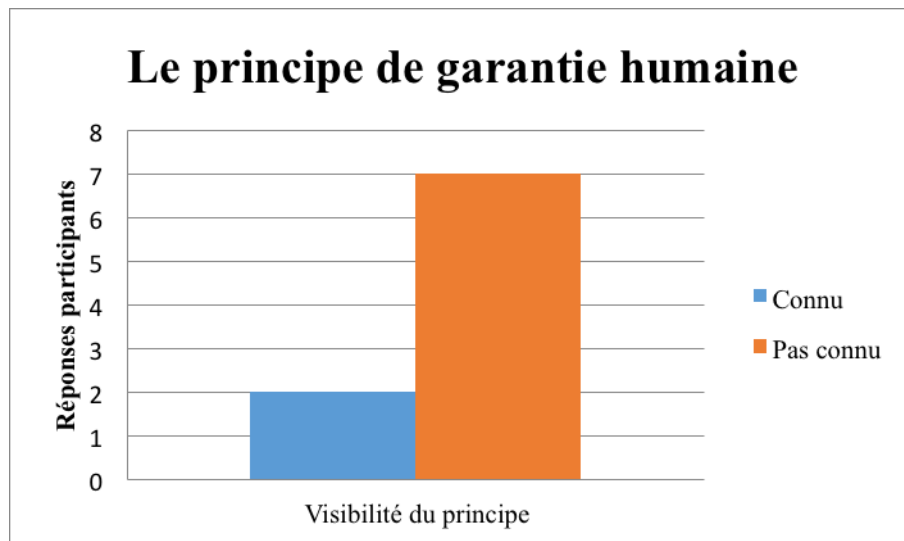
Les personnes ayant accepté de participer à cette recherche étaient des développeurs d'outil d'aide à la décision, des cliniciens, des médecins, des maîtres de conférences universitaires praticien hospitalier, et des informaticiens. Sur les neuf personnes interrogées, quatre participaient au projet de définition d'un algorithme d'IA pour l'AVC du RHU BOOSTER. Trois travaillaient sur la réutilisation de données de santé pour la recherche. Une personne travaillait sur la création des outils d'aide à la décision.

## **4.3 RÉSULTATS THÉMATIQUES**

Les analyses thématiques individuelles ont montré des motifs communs regroupés autour de quatre thèmes principaux :

- La responsabilité ;
- La protection médecin-patient ;
- L'acceptabilité du recours à l'IA dans le parcours de soin ;
- La supervision humaine.

Tableau 2 : Connaissance du principe de garantie humaine



Source : Julie Allouche, 2021

Chaque thème a été divisé en sous-thèmes pour mettre en avant les concepts clefs de la garantie humaine. Un premier constat mérite d'être mis en avant : sur les neuf personnes interrogées, sept ne connaissaient pas le principe de garantie humaine. « *Le terme [de garantie humaine], je ne le connais pas.* » (Entretien 1, 1.163) Six d'entre elles ont formulé une demande de définition de ce principe, bien qu'il leur était rappelé que c'était justement leur conception de ce principe qui était attendu. « *Je me trompe ?* » (E1, 1.159) « *On veut la réponse !* » (E1, 1.170).

#### 4.3.1 Responsabilité

« *À qui l'erreur ?* »

- Entretien 7, ligne 23

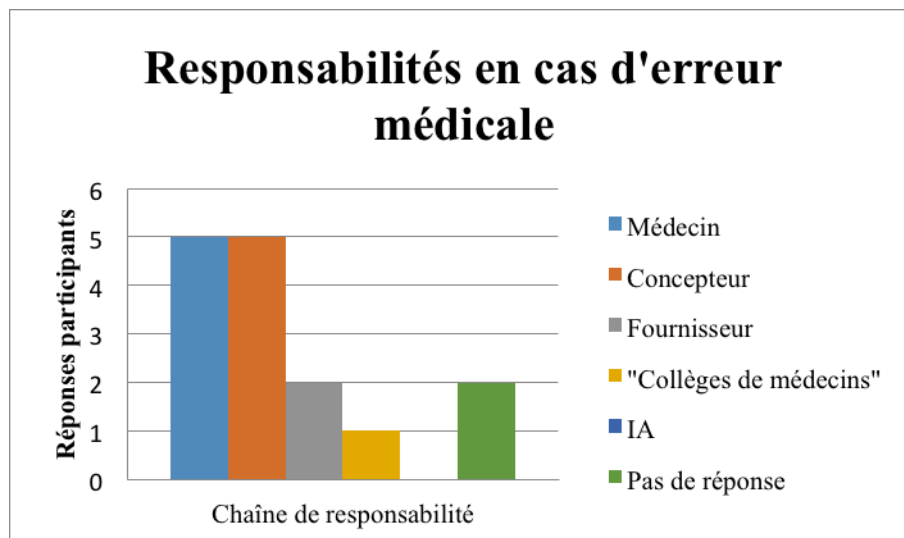
Lors des entretiens, il avait été délibérément choisi de ne pas interroger directement les interlocuteurs sur la question de la responsabilité. Dans notre conception, la question se posait lorsque le médecin avait eu recours à un outil d'IA dans le diagnostic et/ou dans la prise de décision et qu'il commettait une erreur médicale. Fallait-il envisager un rôle de l'IA

dans cette erreur ? Si oui, lequel ? Pour quelles conséquences ? De ce fait, le terme de responsabilité n'a jamais été prononcé en premier de notre part. Le but était de ne pas influencer les participants et d'observer si ce concept avait pour eux, un lien avec le principe de garantie humaine et si oui, lesquels ?

Sur sept entretiens, la question de la responsabilité a été envisagée dans cinq d'entre eux. Pour une personne interrogée, la responsabilité ne fait pas partie de la garantie humaine, elle est plutôt « connexe » à la garantie humaine. C'est donc volontairement qu'elle a écarté ce sujet. Pour le second entretien, la personne n'a jamais mentionné la responsabilité, et ce, sans donner de justification.

Toutefois, dans l'ensemble, les participants ont accordé une place centrale à la responsabilité. « *L'enjeu principal de la garantie humaine repose sur la responsabilité.* » (Entretien 2, ligne 368). C'est une question difficile puisque tous s'accordent à dire qu'elle ne repose pas sur une seule personne. Plusieurs hypothèses sont abordées. « *C'est la leçon de la pluridisciplinarité : un problème a rarement une seule facette. On ne peut pas juste l'aborder sous l'angle de la responsabilité finale. Il va vraiment falloir tout regarder.* » (E1, 1.593). De multiples responsabilités ont été envisagées par les participants, nous allons donc les détailler. En cas d'erreur, la sanction doit être juridique. (E5, 1.229)

Tableau 3 : Types de responsabilité envisagés par les participants



Source : Julie Allouche, 2021

#### 4.3.1.1 Responsabilité du médecin

La responsabilité du médecin est celle qui a été majoritairement évoquée. Pour certains, elle ne fait aucun doute : « *le médecin, il utilise un outil numérique, certes, mais il est toujours responsable de ses actes, de ses décisions.* » (E4, 1.92). « *Dans une décision de soin, je pense que malgré tout, c'est le médecin qui restera responsable, même s'il s'est fait aider ou si sa décision a été assistée par un algorithme, par l'IA.* » (E4, 1.99). En somme, la responsabilité médicale a été justifiée par deux raisons principales ; d'une part, parce que le médecin est libre d'utiliser tel ou tel dispositif d'IA (E4, 1.150 ; E2, 1.169), d'autre, parce que le médecin a des connaissances médicales qui justifient qu'il ne suive pas les recommandations de l'IA lorsque cela est nécessaire. « *Imaginez : je ne comprends pas du tout la machine, mais je comprends la médecine [...] et même si le plus sophistiqué des modèles d'IA me dit « mets le médicament » je ne l'écouterai jamais [...]. La connaissance de base me fait dire que ce n'est bien évidemment pas la bonne décision.* » (E7, 271-287).

D'autres participants ont nuancé la responsabilité médicale. Pour certains, elle est indissociable de l'explicabilité des algorithmes. « *Le médecin est responsable seulement si les informations lui ont été fournies sous une forme compréhensible* » (E1, 1.621). « *Tant au niveau déontologique que réglementaire, on a une obligation médicale, enfin une responsabilité médicale pour expliquer notre décision. C'est pour ça qu'on a besoin de mettre l'accent sur des algorithmes qui nous permettent d'expliquer quelles sont les variables qui leur permettent d'arriver à leur décision. Ne pas juste avoir une boîte noire qui va nous donner un résultat « pronostic bon ou grave » sans qu'on puisse l'expliquer* » (E1, 1.116-120).

Un participant a distingué deux hypothèses pour concevoir la responsabilité médicale (E2) :

1. Quand le logiciel trouve quelque chose et propose un diagnostic, qu'il soit vrai ou faux, le médecin peut vérifier ce diagnostic. Il engage donc à ce titre, sa responsabilité en cas d'erreur, en espérant toutefois qu'il y ait des indicateurs qui permettent de vérifier le diagnostic.
2. Le problème se pose quand le logiciel ne trouve rien alors qu'il y a quelque chose (on parle de « faux négatif ») : que se passera-t-il alors, si le médecin s'est basé sur cette absence de diagnostic pour prendre sa décision ?

#### 4.3.1.2 Responsabilité du concepteur

« *La garantie humaine va être plus au niveau de la conception de l'outil et de la responsabilité des gens qui ont conçu l'outil* » (E2, 1.380). La responsabilité du concepteur est majoritairement envisagée en cas de conception d'un outil sur des données biaisées (E1, 1.629-631), ou plus généralement, en cas de défaillance de ces outils (E4, 1.150).

#### 4.3.1.3 Responsabilité du fournisseur

Le fournisseur n'a été identifié que deux fois comme potentiel responsable. « *Vous pouvez vous retourner contre le fournisseur, mais avec une déclaration de pharmacovigilance.* » (E4, 1.159)

#### 4.3.1.4 Responsabilité des « Collèges de médecins »

Un des participants a considéré que si la responsabilité du médecin peut être engagée, il faut aussi faire peser une responsabilité sur les personnes qui ont autorisé l'utilisation et la mise sur le marché du logiciel défectueux (E7, 1.242-248). Ainsi, en pratique, on pourrait imaginer qu'un « *Collège de médecins* » recommande l'utilisation d'un certain dispositif d'IA pour une situation précise. À l'occasion de la vignette présentée en entretien, un participant a imaginé la mise en œuvre de ce dispositif de la façon suivante : « *un collège de médecins dans les recommandations du SAMU dira : « Pour tous les sites neurologiques du SAMU, on peut appliquer ce logiciel d'IA parce qu'une étude randomisée a montré son utilité »* » (E7, 1.251-257). Or, en pratique, les médecins se basent sur des recommandations de sociétés savantes pour prendre des décisions médicales. Donc, si le médecin se base sur celles-ci, et qu'elles se révèlent être inexactes et le conduisent à une erreur médicale, il pourrait tenter de se dédouaner en établissant un lien de causalité entre la faute et le dommage (E7, 1.257-261).

#### 4.3.1.5 Responsabilité de la machine

La machine peut commettre des erreurs. Comment faire pour rendre ces erreurs moins probables ? « *L'erreur humaine existe aussi, mais la différence est qu'on sait quoi faire pour la rendre moins probable. Que faire si l'IA plante ?* » (E6, 1.28-42). Pour autant, à aucun moment, une responsabilité de la machine n'a été envisagée. « *Vous pouvez pas dire : « c'est pas de ma faute, c'est la machine qui s'est trompée ! »* » (E4, 1.76).

Les participants ont eu beaucoup de difficulté à déterminer à *qui* et *comment* attribuer la responsabilité lorsque l'erreur serait commise par la machine. « *Qui pourrait prendre la responsabilité de la machine ?* » (E6, 1.153).

#### 4.3.1.6 Responsabilité et principe de garantie humaine

L'articulation entre les enjeux de responsabilité et le principe de garantie humaine n'est pas satisfaisante (E7, 1.105). « *Le principe de garantie humaine ne répond pas, pour l'instant, aux enjeux autour de la responsabilité.* » (E7, 1.98). « *Le principe de garantie humaine n'est vraiment pas assez clair par rapport à la responsabilité individuelle du professionnel.* » (E4, 1.127).

Pourtant, les participants ont réclamé à plusieurs reprises une clarification à ce sujet (E4, 1.127) notamment pour éviter « *que ça ne retombe pas finalement à chaque fois sur le petit doctorant pendant sa thèse qui a entraîné un algorithme, revendu ensuite à un énorme acteur de l'intelligence artificielle spécialisée en santé, Google mettons, et que Google se dédouane : « C'est pas moi, c'était Monsieur X. » et donc il y aura plein de questions à se poser* » (E1, 1.633-636).

#### 4.3.2 Protection médecin-patient

« *Le soin de l'être humain ne peut être confié à des machines.* »

- E1, 1.513

La garantie humaine a été, tour à tour envisagée comme un cadre de protection de la relation médecin-patient, comme un cadre de protection du médecin, et comme un cadre de protection du patient. Elle n'a jamais été perçue comme un cadre de limitation au développement des applications de l'IA en santé.

##### 4.3.2.1 La garantie humaine, une source de protection de la relation médecin-patient

« *Pour moi, la garantie humaine c'est déjà : est-ce que la relation médecin-malade est respectée ?* »

- E5, 1.105

« *Le fait qu'il y ait une machine qui s'interpose dans cette relation, est-ce éthiquement acceptable ou déshumanise-t-elle le patient ?* » (E5, 1. 107). La sauvegarde de la relation médecin-patient est souvent mise en avant dans les entretiens. Les participants conçoivent le principe de garantie humaine comme un « gardien » de ce colloque singulier.

Certains craignent une diminution de l'aspect clinique de la pratique médicale. « *Il y aura de moins en moins de médecins au lit du patient.* » (E5, 1.51).

### **Le partage de compétences Homme-machine**

Pour respecter la relation médecin-malade, il faut clairement identifier au préalable quels domaines sont réservés aux humains, et quels domaines peuvent être du ressort des machines. Le social relève du domaine des humains. Certains regrettent qu'en pratique, ce ne soit plus la tendance actuelle. « *Ce que j'aurais aimé un peu plus voir c'est, non des choses qui se prennent pour des humains, mais qu'on réserve l'activité humaine aux humains et l'activité de base de données et de calculs à des machines ; et pas que chacun essaye d'empiéter sur le terrain de l'autre. Aujourd'hui, j'ai l'impression qu'on essaie de faire faire des choses humaines aux machines et des choses de machine aux humains, et ça m'embête* » (E1, 1.548-553).

La place de l'humain, à travers le dialogue avec le patient par exemple, est un élément central de la relation médecin-malade. C'est une relation entre deux humains : « *Que nos outils servent au bénéfice du patient, mais ne soient pas au détriment de la prise en charge humaine du patient. Si on est les meilleurs du monde au diagnostic, mais que finalement, on ne regarde plus jamais les patients, on va rater des trucs complètement objectifs : les contextes de violence sociale ce n'est pas les données qui les diront et ça c'est aussi dans la communication avec le patient que ça se verra. Il y a donc une importance capitale à ce que ça [le dialogue avec le patient] soit en plus et pas en moins* » (E1, 1.399-406).

### **La place du choix dans le recours à l'IA**

Comment articuler la liberté de choix du médecin de recourir à l'IA et l'opposition du patient à ce choix ? « *Si le médecin veut prendre en compte l'algorithme, comment il faudra qu'il justifie ce choix si le patient s'y oppose ?* » (E1, 1.586)

De plus, dans quelle mesure le médecin et le patient peuvent-ils effectivement décider de recourir ou de ne pas recourir à l'IA ? « *Est-ce qu'on peut autoriser un patient à dire : « Moi je ne veux pas que le médecin ait connaissance de l'algorithme. » Auquel cas, comment on fait dans le système pour que le médecin ne le voie pas ?* » (E1, 1.285-287) « *Est-ce que finalement le médecin va pouvoir décider s'il fait confiance ou non à l'algorithme ?* » (E1, 585).



#### 4.3.2.2 La garantie humaine, une source de protection du médecin

« Pour nous [développeurs outils d'IA], la garantie humaine, en gros, c'est le médecin. »

- E2, 1.261

#### **Expérience et formation du médecin**

Un des risques majeurs de l'utilisation du médecin des outils d'IA dans sa pratique quotidienne est une perte d'expertise médicale. Cette perte s'explique par la substitution de tâches jusque là effectuées par le médecin, désormais réalisées par la machine. « *S'il y a plein de choses qui sont faites par l'IA, c'est sûr que les médecins vont être de moins en moins entraînés pour travailler à ce qu'ils font d'habitude. On perdra une grande expertise et ça aussi c'est un problème.* » (E6, 1.48-50 ; en ce sens : E7, 1.223-226 ; E1, 1.522-526)

« [Sur l'utilisation des outils d'aide à la décision utilisant de l'intelligence artificielle] : *la limite c'est que, appliquée en médecine, ça enlève le feeling un peu du médecin. Or, la prise en charge d'un malade ce n'est pas uniquement une décision basée sur des algorithmes.* » (E1, 1.19-22)

« *Il faut changer la manière dont on apprend la médecine aux étudiants.* » (E1, 1.603)

Le corps médical doit être formé aux algorithmes pour comprendre, interpréter les propositions de la machine, et avoir conscience des risques et erreurs possibles de ces utilisations. Le médecin deviendrait en quelque sorte un « *sociologue de la décision* » (E3, 1.247). En pratique, cette formation supplémentaire ne semble pas réaliste, car on exige et on attend déjà énormément des médecins (E1, 1.342-346 ; E1, 1.208-214).

#### **Place du médecin dans la prise de décision finale**

« *La personne qui décide à la fin, c'est le médecin* » (E2, 1.194). Le médecin doit rester le garant de la décision médicale finale (E1, 1.156 ; 166). Deux écueils peuvent entacher cette prise de décision. D'une part, l'explicabilité du résultat par le médecin c'est-à-dire « *que le médecin puisse comprendre pourquoi il a tel résultat* » (E2, 1.263) n'est pas toujours possible (E2, 264). D'autre part, l'impact de la présentation d'une information au

médecin : une même information, présentée de manière différente, peut conduire à une décision différente. « *Une question hyper importante sur la garantie humaine c'est : le chirurgien, il a bien pris une décision à la lecture de quelque chose, mais si on ne lui représente pas l'information de la même façon et qu'il ne prend pas deux fois la même décision, c'est un problème donc est-ce que c'est le chirurgien qui doit être garant ?* » (E1, 455-458).

#### 4.3.2.3 La garantie humaine, une source de protection du patient

« *L'idée de perdre la relation humaine c'est ... (silence)* »

- E6, 1.104

### **Les patients et la téléconsultation**

À titre préliminaire, nous pouvons proposer cette définition de la téléconsultation : « *la téléconsultation est une pratique de télé-médecine et permet à un professionnel médical de consulter un patient à distance.* » (65)

Dans l'échange reproduit ci-dessous, le professionnel de santé interrogé imaginait une situation où le patient ne verrait plus aucun médecin lors de sa téléconsultation.

Un participant a dressé deux portraits de patient type face au recours à la téléconsultation (E5, 1.125-138) :

1. Le patient habitué à la téléconsultation. Il ne sera pas gêné d'avoir son diagnostic sans voir de médecin. Il sera conscient du fait que c'est un médecin qui aura révisé le diagnostic et validé sa prescription.
2. Le patient non habitué à la téléconsultation. Il connaîtra un sentiment de déshumanisation. « *Ce sera une machine qui aura pris la place de l'être humain et ce sera pour lui un ressenti terrible parce qu'il aura l'impression qu'il n'est pas patient finalement, mais qu'il est juste une donnée à analyser. Il perdra cette reconnaissance du patient par le médecin* » (E5, 131-133).

#### 4.3.2.4 Le recours à de l'humain non artificiel

« *Mettre un humain en face d'un humain, c'est encore différent que de mettre un humain face à un écran, puis de mettre un humain face à l'absence d'humain.* »

- E5, 1.134

L'idée que la technologie puisse prendre une place prépondérante dans le parcours de soin d'un patient, au détriment de la relation médecin-patient, inquiète un certain nombre des participants. « [Médecin qui parle] *Si j'étais patient, je ne voudrais pas avoir une machine qui décide.* » (E6, 1.181). Il faut veiller à ce que l'accès aux soins et la qualité des soins des patients soient, au *minimum* équivalents à ce qu'ils ont maintenant, voire supérieurs (E6, 1.127).

Le développement des interfaces artificielles est vécu comme un futur problème pour les patients, dans une approche relationnelle ou encore de niveau de compréhension : « *Pour le patient, je pense que ce n'est pas non plus super évident d'avoir une interface qui n'est pas humaine, et nous [médecins] on travaille avec des personnes âgées, qui ont des fois des problèmes pour s'exprimer, des problèmes de mémoire. Alors déjà c'est difficile de leur proposer une vidéo consultation, donc un petit robot, je sais pas ... [moment d'hésitation]* » (E6, 1.100-103).

#### 4.3.2.5 La protection des données personnelles

La médecine algorithmique repose sur une quantité astronomique de données personnelles. Si son utilisation se développe, ce besoin de données va augmenter. (E3, 1.46) Or ces données, les données de santé par exemple, sont des données sensibles, qu'il faut protéger d'utilisations secondaires pour lesquelles le patient n'aurait pas consenti (E7, 1.82 ; E6, 1.129). Cette protection va dépendre des législations des différents pays (E7, 1.99).

#### 4.3.2.6 Protection et principe de garantie humaine

En matière de protection, la garantie humaine est « *nécessaire, mais insuffisante* » (E1, 1.183). Elle est nécessaire, car elle permet de faire le lien avec le patient, ce qui est capital. Elle est insuffisante, car elle ne garantit pas que le clinicien lui-même ait une compréhension suffisante et éclairée des algorithmes, de leur fonctionnement (E1, 1.216-218).

Avoir un principe juridique fort en matière de protection est devenu un besoin pressant, car les outils d'IA commencent déjà à être conçus. Si la garantie humaine est prévue après leur conception, les développeurs ne pourront pas revenir en arrière dans leur conception (E2, 1.292).

### 4.3.3 Acceptabilité du recours à l'IA en santé

*« Un principe [la garantie humaine] qui permet de désarmer la confrontation entre les médecins et l'IA. »*

- E3, 1.291

Les entretiens ont mis en avant le besoin de rendre plus acceptable une technologie, qui peut sembler nous dépasser à certains moments. Les participants ont expliqué d'où venait ce besoin, et comment le principe de garantie humaine pouvait y répondre. *« Rendre plus acceptable l'évolution qui au fond est écrite, voire prescrite par les technologies aujourd'hui, en l'état de leur développement »* (E3, 1.236).

#### 4.3.3.1 Absence de consensus entre les participants quant à la tournure que va prendre le développement des applications de l'IA en santé

*« L'IA, je ne l'imagine pas vraiment comme une substitution du travail, mais ça va peut-être aller très vite et puis ça va peut-être être beaucoup plus que ce que j'imagine »* (E6, 1.84). Les professionnels de santé interrogés ont du mal à concevoir concrètement les évolutions de l'IA dans leurs pratiques.

Certains se sont montrés favorables à ces évolutions, voire curieux : *« L'IA peut être incroyable. »* (E6, 1.19) ; *« La machine ne remplacera jamais l'humain. »* (E5, 1.63) ; *« J'aimerais bien utiliser l'IA ! Quitte à ne pas lui faire confiance, mais au moins voir comment ça fonctionne. (...) Je suis assez séduite par tout ce qui améliore nos pratiques, tout ce qui est nouveau, tout ce qui est récent. »* (E6, 1.186-191).

D'autres au contraire, sont plus réticents, voire inquiets et craignent la substitution du médecin par la machine : *« L'IA a probablement plus de difficultés que l'humain à donner plusieurs propositions thérapeutiques »* (E6, 1.171) ; *« C'est très difficile pour moi d'imaginer que ces outils soient réellement fiables. »* (E6, 1.91) ; *« Peut-être qu'il y aura une*

*étape intermédiaire, ou peut-être pas, mais j’imagine vraiment plus la substitution* » (E6, 1.91-96).

#### 4.3.3.2 Acceptabilité de ces dispositifs par les médecins, les patients, les décideurs publics et la société

*« Je pense que les médecins, ou les personnes qui travaillent dans le domaine de la santé, ne vont jamais accepter des outils de ce genre s’ils ne sont pas bien régulés, pour protéger leurs patients, leur travail, leur responsabilité médicale et tout ça !* » (UE 6, 1.144-151).

Les difficultés liées à l’acceptation, l’acceptabilité des outils d’aide à la décision, concernent tout le monde : les praticiens eux-mêmes qui pourraient y voir une solution de substitution de la technologie à leur propre expertise, les patients qui pourraient ne pas vouloir se laisser aller entre les mains de technologies mal comprises. Au plan politique, la capacité des décideurs à se doter d’outils prédictifs qui pourraient aussi interférer avec leurs propres capacités à anticiper complique l’acceptabilité de ces dispositifs. En effet, on juge la qualité des politiques à leur capacité d’anticipation. Si les IA sont amenées à interférer avec cette capacité-là, *« est-ce ça pourrait être une difficulté ?* » (E3, 1.46-53).

La France et l’Europe ne sont pour l’instant pas compétitives dans la course technologique de l’IA et des données personnelles. Accepter d’importer des technologies de pays étrangers, c’est perdre en souveraineté, car on ne maîtrise pas la fabrication et les conséquences de ces outils. Le risque de perdre en souveraineté pourrait accélérer l’acceptabilité de ces technologies pour se remettre à niveau au plus vite (E3, 1.85-151).

#### 4.3.3.3 Confiance

La confiance des cliniciens envers les avis automatisés a évolué. Au départ, et pendant longtemps, les cliniciens accordaient peu de confiance à ces avis, souvent par *fatigue d’alerte*. *« Un logiciel de prescription va leur mettre une erreur à chaque fois qu’ils [cliniciens] font une prescription parce qu’il va y avoir une faible interaction entre deux médicaments, ou des choses pas très fondées et donc ils vont finir par cliquer systématiquement par fatigue, parce qu’ils en ont marre, parce qu’ils veulent juste*

*prescrire* » (E1, 1.186- 190). Les cliniciens nourrissaient une certaine défiance envers ces machines. « *Des médecins qui allaient systématiquement à l'inverse de ce que recommandaient les systèmes d'aide à la décision puisque : « Moi je suis humain, je fais la décision, je fais pas confiance à la machine. »* » (E1, 1.192-194). Aujourd'hui, cette défiance persiste, mais les nouvelles générations de médecins sont plus confiantes dans la technologie. Leur masse de travail est énorme et ces outils sont perçus comme une aide. « *On voit qu'il y a de plus en plus de volonté à bien vouloir accepter, voire embrasser ces technologies* » (E1, 1.201).

La confiance des médecins envers les avis automatisés est problématique, car ils engagent leur responsabilité médicale en utilisant ces outils. Certains ont plus de facilité que d'autres à concevoir cette confiance. « *En utilisant ces outils on [cliniciens] se remet quelque part à donner notre confiance à ces outils, en tous cas en y croit suffisamment pour les utiliser et prendre une décision de traitement.* » (E4, 1.28) À l'inverse : « *Si c'est vraiment un outil décisionnel, là je [clinicien] pense que ça serait difficile de lui faire confiance.* » (E6, 1.189) ; « *On fait bêtement confiance à la machine et on en oublie les bases de la médecine.* » (E7, 1.215).

D'autres s'interrogent sur la fiabilité de ces outils. « *À quel moment on décide qu'un logiciel est suffisamment fiable pour ne plus avoir de double tchek avec un médecin derrière qui valide le diagnostic ?* » (E6, 1.338-343).

#### 4.3.3.4 Cadre légal

« *[En parlant des outils d'IA] ça pose quand même beaucoup de questions et on n'a pas encore toutes les réponses.* »

- E7, 1.26

« *Pour utiliser l'IA, il faut montrer qu'il y a un réel intérêt : un intérêt scientifique, un intérêt financier. Il faut que ça réponde à des principes de garantie, de protection des données et de responsabilité. Si tout ça c'est cadré, il n'y a aucun problème à son application à mes yeux, mais on en est encore loin encore, je pense* » (E7, 1.113-116). Un cadre légal précis en matière d'IA en santé est attendu des participants.

Cet encadrement doit se faire avant que l’outil d’IA se développe ou en même temps, mais surtout pas après. « *Après ça serait une catastrophe, des dérives monumentales* » (E7, 125 ; 1.132-135 ; 1.128-130).

Ce cadre légal est préférablement attendu au niveau européen (E7, 1.145-151). Bien que les participants réclament un cadre légal, ils reconnaissent méconnaître l’état de la législation actuelle à ce sujet-là (E7, 1.212 ; E1, 1.619).

#### **4.3.4 Supervision humaine**

*« L’IA c’est l’assistance de l’Homme par la machine ; et la Garantie Humaine c’est l’assistance de la machine par l’Homme, en s’assurant que la machine ne s’est pas plantée, qu’il y ait une relecture, une vigilance. »*

– E3, 1.161

La garantie humaine est la garantie d’une intervention humaine de l’ordre de la supervision de l’outil d’IA pendant sa conception et pendant son utilisation.

##### **4.3.4.1 Supervision humaine au stade de la conception de l’outil d’IA**

*« Une machine construite par des gens qui ne sont pas du métier est une machine qui sera sûrement une source d’erreur et ne sera presque pas éthique. »*

- E5, 1.93

#### **Reproduction des biais cognitifs et recherche d’équité**

À titre préliminaire, un rappel de la définition des biais cognitifs : « *Une erreur cognitive est un mécanisme inconscient de la pensée à l’origine d’une décision erronée. Lorsque la déviation de la pensée logique par rapport à la réalité devient systématique, on parle de biais cognitifs* » (2).

Plusieurs concepteurs et professionnels de santé interrogés alertent sur la reproduction des biais cognitifs dans la conception des outils d’IA. « *À partir du moment on est sur des méthodes d’apprentissage, si on a un biais dans les données, quel qu’il soit, on va pouvoir*

*produire des méthodes prédictives biaisées* » (E1, 1.123). Le risque est que l'outil soit utilisé de manière indifférenciée sur la population générale. Or, alors il est conçu sur des données à propos d'une partie bien spécifique de la population et ne peut donc pas être généralisé. En effet, la population spécifique ne s'adapte pas forcément au reste de la population.

Ainsi, les participants ont établi un lien entre biais cognitifs et équité. « *Il y a des données qui ne sont probablement pas pertinentes à prendre en compte et qui sont quand même prises en compte dans ces modèles et qui conduisent à d'énormes problèmes de validité externe.* » (E1, 1.136-138). « *On a des traitements qu'on utilise pour tout le monde, mais on s'aperçoit que, et on le sait, on est tous différents. On ne réagit pas tous de la même façon aux traitements.* » (E4, 32).

« *Sur la question d'équité, l'outil doit marcher pour tout le monde, qu'il n'y ait pas de laisser pour compte.* » (E2, 1.381) Il faut être clair si l'outil a été conçu pour fonctionner sur une population générale ou spécifique (E2, 1.342).

La validité de la machine dépend donc de la qualité des informations : « *La machine va pouvoir, à condition que je lui ai donné les bonnes informations, analyser ces bonnes informations qui relèvent de la pratique et de l'expertise du métier.* » (E5, 1.91).

## **Une équipe hétérogène**

Un des critères récurrents de performance des outils d'IA est la participation des professionnels de santé au stade de conception de l'outil, et tout au long de l'utilisation en vie réelle de ces outils. « *Sur la phase de conception, nous [développeurs d'outil d'aide à la décision] on travaille de toute manière en partenariat avec les médecins en permanence quand on crée ces outils-là. [...] Ça permet d'éviter de faire des absurdités ou de ne pas prendre en compte des facteurs auxquels on n'avait pas pensé, alors que c'est des choses évidentes pour les médecins qui vont les utiliser* » (E2, 1.265-270). Un outil d'aide à la décision construit sans l'aide des professionnels de santé, serait un outil qui ne répondra pas à la demande et à l'attente, et qui sera source d'erreurs (E5, 1.41-43). Il est nécessaire d'intégrer les médecins, le personnel médical et paramédical au processus de conception de ces outils. (E7, 1.60-66).

À ce propos, un des participants a proposé la création d'un « *Collège de médecins* » qui validerait en amont de l'utilisation, l'efficacité et l'intérêt médical de l'outil (E7, 1.245-247).



#### 4.3.4.2 Supervision humaine au fil de l'utilisation pratique de l'outil d'IA

Pour que l'outil soit toujours aussi performant, il ne suffit pas de prévoir une supervision humaine au stade de la conception, car les résultats obtenus pendant les phases d'essais ne sont pas toujours ceux obtenus en vie réelle. Ainsi, une supervision humaine doit être aussi envisagée tout au long de son utilisation. Un contrôle régulier de la capacité des algorithmes à être tout le temps aussi performant doit mis en place (activité de veille). De même, il faut aussi mettre en place un contrôle au moment du résultat attendu. Enfin, il faut s'assurer que la prescription soit toujours médicale. « *La machine peut me proposer le diagnostic, mais c'est au médecin de valider le diagnostic réalisé par la machine* » (E5, 1.96 ; dans le même sens : E1, 1.156 ; 165-167).

La surveillance des outils devrait permettre un retour sur tous les outils algorithmiques afin de les améliorer constamment. Le patient doit être intégré à cette amélioration pour que ces outils soient satisfaisants (E6, 1.132).

Un participant avait proposé la création d'une unité de référence, dans chaque groupe hospitalier, qui « *soit à disposition, qui surveille de loin tout ça, et qu'on puisse contacter si on a un doute ou autre, et qu'il fasse des révisions programmées des contrôles avec de faux patients par exemple, régulièrement, sur tous les systèmes* » (E6, 1.259-262).

## 5 : DISCUSSION

---

*« L'indépendance n'est pas un privilège du médecin, mais un droit indiscutable du malade. C'est ce caractère un peu sacré de l'existence de la médecine, c'est cette responsabilité directe que le médecin ne peut ni refuser ni partager qui réclament une indépendance que bien peu d'autres professions sont en droit de réclamer. »*

- Président du Conseil de l'Ordre, 1959 (66)

### 5.1 HYPOTHÈSES DE RECHERCHE ET RÉSULTATS

Deux hypothèses ont guidé notre recherche depuis son commencement. Pour rappel, les hypothèses de départ étaient les suivantes :

1. La première hypothèse était celle d'une notion peu connue du principe de garantie humaine du numérique en santé, due à son caractère nouveau et à la complexité des évolutions législatives.
2. La seconde hypothèse avait trait à la responsabilité : les concepteurs d'outils d'aide à la décision ne se considèrent actuellement pas responsables en cas d'erreur médicale et la faute professionnelle est imputée aux soignants.

Désormais, il convient de les confronter aux résultats de la recherche pour voir si ces derniers confirment, nuancent ou passent sous silence les hypothèses de départ.

Concernant la première hypothèse, un bref sondage des résultats a permis d'établir que la majorité des participants à la recherche méconnaissaient le principe de garantie humaine du numérique en santé (*cf.* Tableau 2). Seulement deux participants sur neuf avaient connaissance de ce principe. Cette méconnaissance était prévisible, car le principe de garantie humaine est nouveau et son cadre juridique n'est à ce jour pas encore fixé. Néanmoins, deux participants avaient connaissance du principe de garantie humaine. Le premier, car il échangeait fréquemment avec Monsieur David Gruson, défenseur de la notion de garantie humaine. Le second, car il disait avoir « suivi le débat » (E3, 1.156), sans donner plus de précision sur la nature de ce débat.

Toutefois, cette méconnaissance du principe ne fut pas un obstacle aux échanges avec les participants, car ils avaient connaissance des enjeux autour du développement des applications de l'IA en santé. C'est d'ailleurs ces enjeux, tels que la responsabilité, la protection médecin-patient, la capacité à rendre plus acceptable l'IA en santé auprès des professionnels de santé, des patients, des décideurs politiques et de la société qui sont ressortis de l'analyse des entretiens. La réelle question sous-jacente est de savoir si le principe de garantie humaine répond à ces enjeux, et de quelles façons. Nous tenterons de répondre à cette question tout au long de la discussion.

Outre le caractère nouveau de ce principe, une autre explication est apparue dans les résultats pour corroborer cette hypothèse. Un « mur » semble séparer la médecine du droit. « *Je comprends jamais quand le droit parle* » (E1, 1.619). La technicité des textes juridiques et le manque de visibilité des activités du Parlement dans les médias depuis les États généraux de la bioéthique en 2018 sont peut-être à l'origine du manque d'interactions entre ces deux pôles. Ce constat est à notre sens regrettable, car ils sont tous deux interdépendants. L'exercice de la médecine est en partie régulé et socialement acceptable grâce à la Loi (on peut facilement le concevoir après les atrocités de la Seconde Guerre mondiale, pour ne citer qu'un exemple) ; de même que le droit n'a de sens que pour ce qu'il apporte aux domaines extérieurs à lui-même, tels que la médecine.

Paradoxalement, si les participants reconnaissent ne pas être informés de l'existence du principe de garantie humaine, ou encore de l'état de la législation en matière d'intelligence artificielle et de protection des données de santé, ils réclament unanimement un cadrage juridique en la matière, et ce au plus vite. « *Heureusement qu'on se pose la question légalement pour encadrer ça le plus tôt possible avant de se retrouver face à ce qui est déjà fait* » (E1, 1.657). Cette demande est tout à fait légitime, car les concepteurs et *a fortiori* les professionnels de santé sont les premiers concernés par le développement des applications de l'IA en santé. Les professionnels de santé utilisent dans leur pratique quotidienne les outils conçus par les concepteurs. Les patients sont tout aussi concernés, mais à défaut de les avoir intégrés à l'échantillonnage, nous ne pourrions pas nous prononcer sur leur position respectivement à ce sujet.

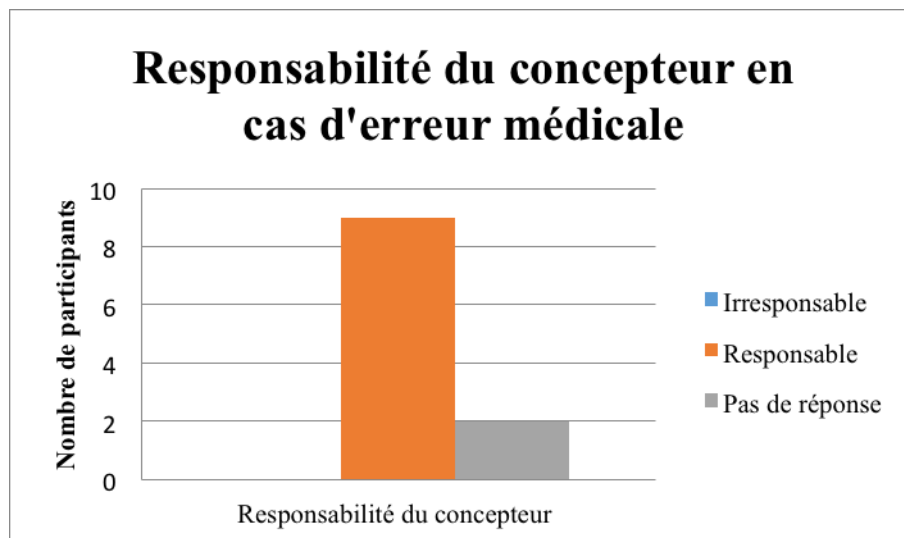
Les regards neufs que portaient les concepteurs et les professionnels interrogés sur le contenu du principe de garantie humaine étaient complémentaires. Certains concevaient la garantie humaine comme un cadre de protection exclusive du médecin ; d'autres, exclusivement du point de vue du patient. Enfin, certains englobaient le médecin et le patient dans l'expression de garantie humaine. Ils s'accordent à dire qu'elle est nécessaire, mais insuffisante. Ils en attendent plus de ce principe, car les enjeux du développement de l'IA en santé sont très nombreux et pourraient avoir des conséquences majeures sur les pratiques de soin. Le principe de garantie humaine a-t-il vocation à garantir la place essentielle de l'humain au cœur du soin ? Si oui, comment ?

Concernant la deuxième hypothèse, la question de la responsabilité a été spontanément abordée par sept des neuf participants. C'est un thème récurrent des entretiens qui préoccupent les personnes interrogées. Cette seconde hypothèse regroupe deux idées :

1. En cas d'erreur médicale, les concepteurs ne conçoivent pas leur responsabilité.
2. En cas d'erreur médicale, les concepteurs attribuent la responsabilité aux soignants.

La première idée peut être étudiée à l'aide du graphique suivant :

Tableau 4 : Responsabilité du concepteur en cas d'erreur médicale



Source : Julie Allouche, 2021

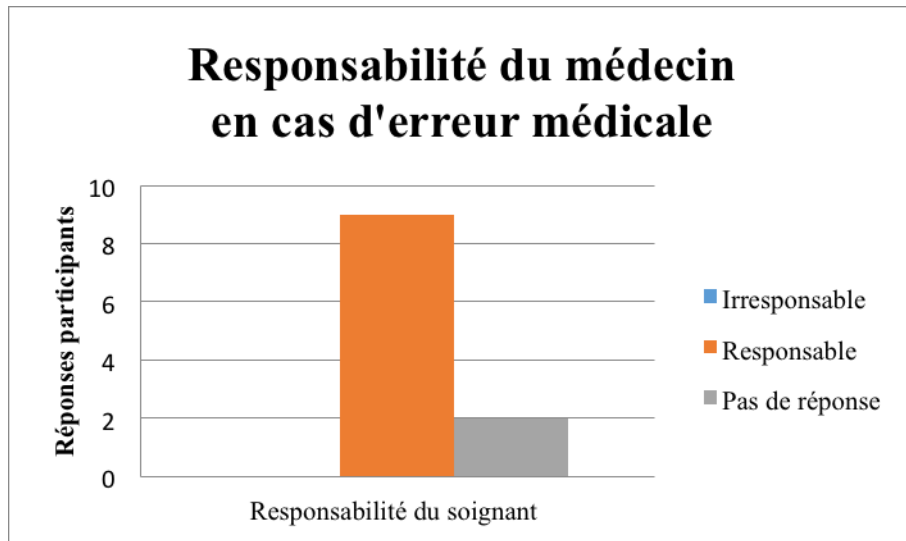
Ce graphique ne confirme pas cette idée. En effet, sur les sept personnes interrogées qui ont abordé la question de la responsabilité, sept envisageaient la responsabilité du

concepteur en cas d'erreur médicale lorsque le professionnel de santé avait utilisé un outil d'aide à la décision intégrant de l'IA. Dans ce cas-là, la responsabilité du concepteur est envisagée par les concepteurs eux-mêmes en cas de défaillances des dispositifs (E4, l.150).

Aucun des sept participants n'a réfuté la responsabilité du concepteur ou ne l'a pas abordée. Ce constat est pertinent en ce qu'il démontre la conscience des concepteurs interrogés sur les dérives potentielles des outils qu'ils développent. On peut comprendre qu'ils aient besoin d'un cadre juridique clair et solide pour orienter leur construction. En effet, ils ne pourront pas revenir, ou difficilement semble-t-il, sur leurs outils une fois construits. Une des solutions proposées pour atténuer ces erreurs est d'intégrer dès la conception de l'outil d'aide à la décision ou au diagnostic, les professionnels de santé spécialisés dans la spécialité médicale où l'outil sera utilisé. Par exemple, si un outil d'aide à la décision est conçu pour assister les médecins en cas d'AVC, il faudrait intégrer des neurologues au développement de ces outils pour s'assurer qu'ils répondent de façon adaptée aux besoins des médecins en situation réelle. Les hypothèses de défaillances des dispositifs ont peu été développées. Il est donc difficile de proposer d'autres solutions sans connaître la nature de ces défaillances.

L'idée selon laquelle le principe de garantie humaine du numérique en santé a vocation à régir les enjeux autour de la responsabilité n'est pas unanime. Une position minoritaire d'un professionnel de santé conçoit la responsabilité de façon connexe à la garantie humaine. Il ne l'intègre pas dans la définition de la garantie humaine. Cette réflexion fera l'objet d'une analyse plus précise dans la seconde partie de la discussion (*cf.* 5.2 Confrontations avec la littérature). Un autre professionnel de santé n'a pas abordé la question de la responsabilité. Afin de ne pas l'influencer, la question ne lui a pas été posée. La responsabilité n'entre donc pas de façon évidente et systématique dans les représentations de la garantie humaine. Il faudra donc s'interroger sur la pertinence de la responsabilité comme composante de la garantie humaine.

Tableau 5 : Responsabilité du médecin en cas d'erreur médicale



Source : Julie Allouche, 2021

En cas d'erreur médicale, le médecin, bien qu'il ait eu recours à un outil d'aide à la décision utilisant de l'IA dans son processus de décision, engage sa responsabilité. Cette réponse est sans appel pour les participants. Cette hypothèse est corroborée par les résultats de la recherche.

Toutefois, quelques nuances sont nécessaires à ce propos. Responsabilité médicale oui, mais sous réserve de :

- la qualité de l'information : si le médecin a pris une décision erronée sur la base de données biaisées, est-il toujours responsable ? La réponse n'est pas évidente. Dans cette situation, peut-être faudrait-il établir si, grâce à ses connaissances médicales et sa pratique, le médecin était en mesure de constater l'erreur. Les outils d'« aide » à la décision sont-ils vraiment une aide pour le médecin si celui-ci doit en permanence se méfier des résultats proposés par l'outil ? Comment faire confiance à ce qui peut nous tromper ? À ce sujet, un professeur des universités-praticien hospitalier (PU-PH) interrogé répondait que le risque d'erreur pouvait se compenser par la performance de l'outil. Autrement dit, si l'outil est performant dans 98% des cas, et qu'il commet des erreurs graves dans 2% des cas, il faut continuer à l'utiliser, car le service rendu pour les autres patients est très important. Comment s'assurer que la prédiction et la recommandation fournies par les algorithmes ne soient bien qu'une aide à la prise de décision et à l'action humaine, sans aboutir à une déresponsabilisation et une perte d'autonomie ?

- l'explicabilité du résultat au médecin : le médecin a-t-il besoin de comprendre le résultat de la machine pour prendre sa décision ? Certains concepteurs et professionnels de santé ont mis en avant cette difficulté. Toutefois, le médecin d'aujourd'hui ne comprend pas le fonctionnement interne de tous les outils technologiques qui lui permettent de prendre sa décision ; et cela ne l'empêche pas de la prendre pour autant. Par exemple, en cas d'AVC, le médecin va procéder à une imagerie pour déterminer le type d'AVC et traiter son patient selon ce résultat. Il ne maîtrise pas le fonctionnement interne d'une imagerie et pourtant il fait confiance au résultat de l'imagerie. Peut-on appliquer ce raisonnement à l'IA ? Peut-être la capacité autoapprenante de certaines IA justifie-t-elle de se poser la question. C'est cette capacité d'autoapprentissage, dite « *machine learning* », qui la distingue des autres technologies utilisées aujourd'hui par le médecin dans le diagnostic du patient et le choix du traitement. De plus, l'intérêt de comprendre le fonctionnement interne d'un outil technologique est d'être en mesure de le rectifier *a posteriori* en cas d'erreur. Ce n'est pas au médecin de le faire bien sûr, mais qu'en est-il des concepteurs ? Seront-ils capables de rectifier l'IA ?

Une dernière remarque sur cette seconde hypothèse de recherche en matière de responsabilité : elle est doublement incomplète. Elle n'envisage que la responsabilité du concepteur et celle du professionnel de santé. D'une part, les résultats de la recherche sont venus compléter la chaîne de responsabilité en concevant les responsabilités de l'hôpital, du fournisseur, du producteur, du testeur et de la personne assurant la maintenance de l'outil d'IA en cas d'erreur médicale. Faut-il les intégrer au dispositif de responsabilité ? Si oui, comment ? Qui d'autre pourrait être tenu responsable ? D'autre part, les résultats de la recherche ne sont pas favorables à une responsabilité de l'IA. Il est vrai que derrière chaque machine, il y a des humains qui ont construit ces machines. En dépit de cette présence humaine à l'origine de la création des outils, les IA autoapprenantes évoluent d'elles-mêmes au fur et à mesure de leur utilisation. Pourrait-on considérer qu'à partir d'un certain stade d'autoapprentissage, elles s'éloignent tant de leurs capacités d'origine que les concepteurs ne pourraient plus être tenus pour responsables en cas de dysfonctionnement ? En d'autres termes, jusqu'à quand le concepteur d'une IA autoapprenante est-il réellement son concepteur si le fonctionnement interne de cette IA venait à lui échapper ?

## 5.2 COMPARAISON AVEC LA LITTÉRATURE

Les résultats de l'étude ont permis de mettre en avant quatre thèmes en lien avec le principe de garantie humaine :

- La responsabilité ;
- La protection médecin-patient ;
- L'acceptabilité du recours à l'IA dans le parcours de soin ;
- La supervision humaine.

Nous allons dès à présent discuter ces résultats à la lumière de la littérature sur le principe de garantie humaine.

### 5.2.1 Responsabilité

Comment expliquer que la responsabilité soit quasiment abordée dans tous les entretiens et qu'elle ne figure dans aucune version de l'article 11 du projet de loi relatif à la bioéthique ? Un consensus semble émerger du côté des législateurs pour expliquer cette absence : les régimes juridiques actuels de responsabilité des professionnels de santé, du fait des choses et des produits défectueux suffisent à couvrir les hypothèses de responsabilité en matière d'IA en santé. Si cela est vrai pour le présent, et encore faudrait-il le démontrer, jusqu'à quand nos régimes de responsabilité actuels seront cohérents avec l'avancée si rapide des technologies ? Dans les prochaines années, comment adapterons-nous notre arsenal législatif en matière de responsabilité et d'IA ? Faudra-t-il privilégier le niveau interne ou européen ? La responsabilité d'un constructeur de voitures d'IA est-elle équivalente à celle d'un concepteur d'outil d'IA ? Les États généraux de la bioéthique ont montré qu'il était difficile de comprendre la portée des innovations scientifiques et technologiques, car elles sont complexes et le niveau général de connaissance est insuffisant (9).

L'explicabilité des algorithmes est-elle le corollaire de la responsabilité médicale ? Une forte convergence des participants tend à répondre favorablement à cette question. *« Tant au niveau déontologique que réglementaire, on a une obligation médicale, enfin une responsabilité médicale pour expliquer notre décision. C'est pour ça qu'on a besoin de mettre l'accent sur des algorithmes qui nous permettent d'expliquer quelles sont les variables*



*qui leur permettent d'arriver à leur décision. Ne pas juste avoir une boîte noire qui va nous donner un résultat « pronostic bon ou grave » sans qu'on puisse l'expliquer » (E1, 1.116-120).*

Néanmoins, une position différente peut être construite. Lors de nos échanges, Monsieur David Gruson nous met en garde contre le caractère séduisant du concept d'explicabilité des IA autoapprenantes pour trois raisons. Premièrement, parce qu'il faudrait passer un temps incroyable, à supposer que ce soit même possible, pour lister tout ce que fait la machine à chacune de ses étapes de raisonnement. De plus, ce qui est explicable n'est pas forcément juste d'un point de vue éthique. Par exemple, si une IA pendant un risque épidémique venait à protéger la population au prix du sacrifice de groupes d'individus ; on ne considérerait pas forcément cela juste du point de vue de l'éthique. Par contre, on peut parfaitement expliquer le rationnel de cette décision. La troisième objection à l'explicabilité comme norme juridique tient à l'essence de la médecine : la médecine avance par des sauts qui au départ sont en partie inexplicables. Si on avait imposé le principe d'explicabilité à Pasteur, il n'aurait sans doute pas pu inventer ses vaccins (44). Le principe d'explicabilité des algorithmes fait donc l'objet d'un désaccord raisonnable.

Tant les participants à l'étude, que les juristes, s'opposent à une responsabilité de l'IA, car derrière toute machine, il y a des humains (dans la conception, dans la prise de décision, etc.). Certes, mais qu'en est-il des machines autoapprenantes ? Pourrions-nous considérer qu'à partir d'un certain niveau, la machine se serait tant éloignée de sa conception initiale, que le concepteur ne pourrait plus être tenu responsable ? *Quid* si l'IA venait à être capable de créer elle-même d'autres IA ? Comment garantir qu'un patient victime de l'erreur d'un algorithme puisse être indemnisé de son préjudice ? Le Parlement européen a envisagé en 2017 d'attribuer une personnalité juridique aux algorithmes dotés de capacités d'autoapprentissage (67). Le CCNE, le Conseil d'État, et la majorité des experts auditionnés par la Mission d'information sur la révision de la loi bioéthique y sont défavorables (38).

### **5.2.2 Protection médecin-patient**

Nous retrouvons dans les résultats et dans les versions 2 et 3 de l'article 11 du projet de loi bioéthique, la faculté du médecin de recourir à l'IA. Dans quelle mesure la liberté de choix d'utiliser l'IA ou de ne pas l'utiliser, d'utiliser un dispositif d'IA plutôt qu'un autre est-

elle effective dans des situations médicales d'urgence ? Des pressions exercées sur le médecin pourraient relativiser cette liberté. Si le patient venait à pouvoir refuser l'utilisation de l'IA, quelle liberté resterait-il au médecin ? Faut-il prévoir un droit de non-recours du médecin à l'IA et un droit d'opposition du patient à l'utilisation d'un TADM dans son parcours de soin ? Comment articuler ces deux droits ?

En ce sens, l'article 14 du projet de règlement sur l'intelligence artificielle de la Commission européenne ajoute un élément nouveau par rapport à la rédaction de l'article 11 du projet de loi de bioéthique, qui est une capacité d'interruption de l'IA à tout moment. Pouvons-nous réellement interrompre l'IA à tout moment ? Comment s'assurer que cette faculté ne soit pas dénuée de sens en pratique ?

La maîtrise de la décision finale par le professionnel de santé est envisagée par le législateur et par les participants. Toutefois, comment prévenir le risque que cette maîtrise ne soit qu'une façade si le médecin venait à perdre trop d'expérience médicale à cause d'une utilisation massive de ces outils, couplée d'une absence de formation à leur usage (cristallisation du savoir) ? Comment s'assurer que la prédiction et la recommandation fournies par les algorithmes ne soient bien qu'une aide à la prise de décision et à l'action humaine, sans aboutir à une déresponsabilisation et une perte d'autonomie ? Le médecin est-il le véritable décideur ou le simple auxiliaire de l'algorithme ? À partir de quand considère-t-on que le médecin est suffisamment intervenu dans le processus de décision ? Comment se prémunir du simple « oui » ?

Les résultats de l'étude corroborent deux risques éthiques mis en avant par le CCNE dans l'avis 129 (15). D'une part, le patient risque de perdre en autonomie dans sa prise en charge. Faut-il qu'il ait accès au résultat fourni par l'algorithme ou seulement à l'appréciation formulée par le médecin ? (38) D'autre part, les outils conçus sur une population spécifique et utilisés sur une population générale ne seraient pas adaptés à celle-ci. « *La souffrance est privée, mais la santé est publique* », disait Paul Ricœur. Cette citation illustre la dualité des intérêts de l'individu VS du collectif. Comment garantir que le collectif n'écrase pas les singularités et les libertés individuelles ? (5)

Le besoin de renforcer le régime juridique des données de santé, notamment les données de santé collectées hors du cadre du soin, est confirmé par les résultats de l'étude. Si ce besoin se confirme, la réponse n'est pas évidente : s'il est vrai que l'Union européenne, à

travers le RGPD, tente de réguler la protection des données personnelles ; l'enjeu est de taille, car ce sont les GAFAM (Google, Apple, Facebook et Microsoft) qui détiennent majoritairement nos données, et non d'autres États. La difficulté repose sur la perte de souveraineté des États en la matière. Comment y remédier ? Comment veiller à ce que l'exploitation des données ne se substitue pas à la décision humaine dans le soin et que des garde-fous soient mis en place pour protéger la vie privée ? (5)

D'où vient ce besoin, répété dans les résultats et dans la littérature, de protéger le médecin et le patient de l'IA ? Pourquoi garantir la maîtrise de la décision finale au médecin si l'IA se révèle être plus performante que lui ?(21) Pourquoi est-ce si difficile de (re)mettre de l'humain au cœur du soin ? N'est-ce pas avant tout la dimension humaine qui fait la richesse du soin ? N'est-ce pas l'humain qui donne le sens profond du soin et qui soutient et anime les soignants, quelles que soient les difficultés ?

Nous pensons que mesurer quantitativement le bénéfice des relations humaines est si difficile, qu'on a tendance à l'oublier. Ci-dessous (figure 7) le dessin d'un enfant qui représente un patient assis sur le lit d'auscultation, les parents assis sur les chaises dans le bureau du médecin et le médecin dos à tout le monde, entrain de taper sur l'ordinateur. Ce dessin semble illustrer la perte de lien humain dans la relation de soin à cause de l'utilisation massive de la technologie. Quel est le sens du soin ?

Figure 7 : Un dessin d'enfant représentant la consultation avec le médecin à l'ère des ordinateurs portables dans la salle d'examen



© 2011 Thomas G. Murphy, MD.

Source: Thomas G., Murphy, MD, *The Cost of Technologie*, 20/06/2012 (68)

### 5.2.3 Acceptabilité du recours à l'IA en santé

Nous avons été surpris de l'importance accordée par les participants à l'acceptabilité de l'usage de ces outils. À vrai dire, c'est un thème que nous n'avions pas envisagé au départ et qui est peu ressorti de notre étude de la littérature. Pourquoi le recours à l'IA en santé ne fait-il pas l'unanimité ? Que faut-il vraiment rendre acceptable : le recours à l'IA en lui-même ou les conséquences de ce recours ? Quelles sont les craintes de ses opposants ? Sont-elles fondées ? Vont-elles être prises en compte dans la régulation de l'IA ? Rendre acceptable une technologie par un cadre juridique est-elle la réponse adaptée ?

Certains participants craignent qu'à terme, l'IA se substitue au médecin. L'Assemblée nationale rappelle à ce sujet que « *les algorithmes sont dénués de toute capacité*

*de communication et d'empathie et ne savent pas prendre en compte le patient dans son intégralité. Les algorithmes ne peuvent être qu'un complément de l'activité du médecin sans jamais s'y substituer* » (38). Cette position de l'Assemblée nationale sera-t-elle identique lorsque les algorithmes seront dotés de ces capacités et auront accès à toujours plus de données du patient ?

#### **5.2.4 Supervision humaine**

Trois types de contrôle sont ressortis de l'analyse des résultats et de l'étude de la littérature:

- Une supervision humaine au stade de conception de l'outil ;
- Une supervision humaine au stade de la délivrance du résultat de la machine ;
- Un contrôle humain régulier tout au long de l'utilisation de l'outil.

La proposition de l'un des participants de créer une unité de référence dans chaque groupe hospitalier qui surveille l'utilisation des outils d'IA par exemple grâce à des révisions programmées, des contrôles avec de "faux patients" sur tous les systèmes (E5, 1.258-262) ressemble à la proposition des "collèges de garantie humaine" de Monsieur Gruson (41).

Nous pensons que la supervision humaine est ambiguë dans la quatrième version de l'article 11 du projet de bioéthique.

Quelle est l'utilité de l'algorithme si l'Homme repasse derrière ? L'Homme contrôle-t-il l'outil ou l'outil contrôle-t-il l'Homme ? (69)

### **5.3 LIMITES ET FORCES**

#### **5.3.1 Limites de l'étude**

Les résultats de cette analyse d'entretiens, bien qu'intéressants, posent la question des limites et biais de l'étude.

Nous pensons que l'une des limites principales de la méthode employée au sein de cette recherche repose sur la nature majoritairement juridique ou à « vocation juridique » des sources de la bibliographie, par exemple les textes préparatoires qui ne sont pas des textes de nature juridique, mais qui s'inscrivent dans le processus de révision d'une loi. En effet,

l'arrière-plan théorique et conceptuel a été construit dans l'ensemble, sur l'étude des textes préparatoires à la révision de la loi de bioéthique et aux quatre versions de l'article 11 du projet relatif à la révision de la loi de bioéthique. Or les résultats de l'étude, de par la formation médicale et/en informatique des participants à l'étude, ne se basent pas sur des connaissances juridiques, bien au contraire. Ceci a certainement favorisé un décalage de compréhension du sujet entre le recruteur et les participants.

Nous regrettons, faute de temps, de ne pas avoir pu pallier ce biais par les nombreuses recommandations de lectures scientifiques, sociologiques et philosophiques de nos professeurs, intervenants au Master et participants à l'étude.

Cette préoccupation rejoint d'ailleurs la deuxième limite que nous envisageons, à savoir que notre seconde hypothèse de départ est sans doute mal formulée et imprécise. L'hypothèse en question était la suivante : les concepteurs d'outils d'aide à la décision ne se considèrent actuellement pas responsables en cas d'erreur médicale et la faute professionnelle est imputée aux soignants. Deux critiques à cet égard. Premièrement, nous aurions dû préciser le type d'erreur médicale auquel nous faisons référence pour comprendre le lien avec la responsabilité du concepteur. Une formulation plus juste aurait pu être celle-ci : les concepteurs d'outils d'aide à la décision ne se considèrent actuellement pas responsables en cas d'erreur médicale lorsque le médecin réfèrent a recouru à un outil d'aide à la décision intégrant de l'IA ; et la faute professionnelle est imputée aux soignants.

Deuxièmement, le terme de soignant est trop vague et interroge sur la responsabilité du reste de l'équipe médicale. Or, la responsabilité pèse sur le médecin qui a pris la décision finale et non pas sur les soignants en général. C'est d'ailleurs ce qui justifie l'autonomie du médecin à travers sa liberté de prescription et sa liberté de jugement. Une formulation plus juste aurait pu être celle-ci : les concepteurs d'outils d'aide à la décision ne se considèrent actuellement pas responsables en cas d'erreur médicale lorsque le médecin réfèrent a recouru à un outil d'aide à la décision intégrant de l'IA ; et la faute professionnelle est imputée à ce dernier.

Les résultats de l'étude révèlent les positions de neuf professionnels, dont huit issus de trois centres français concernés par la recherche. Ils n'ont donc pas vocation à tenir lieu de conclusions généralisables, ce d'autant qu'ils reposent sur une analyse qualitative. Un biais potentiel de sélection tient au fait qu'il s'agit de professionnels ayant accepté de participer à l'étude et dont les motivations variées peuvent avoir influencé les résultats. Deux d'entre eux

étaient par ailleurs intervenus dans les unités d'enseignement de notre Master. Nous avons donc déjà eu un premier contact avant l'entretien. Il est possible que, sur des sujets interrogeant leurs pratiques professionnelles, le caractère semi-directif des entretiens, ainsi que l'enregistrement de ceux-ci, malgré l'engagement d'anonymisation, ont pu être la source d'une autocensure.

Quatre des neuf participants travaillent ensemble au projet de définition d'un algorithme d'IA pour l'AVC (RHU BOOSTER). La question d'une politique commune à l'ensemble de ces participants a peut-être influencé leur regard sur les enjeux de l'IA en santé favorisant une approche commune de ces enjeux.

Pour des raisons d'organisation et de confidentialité, nous n'avons finalement pas pu assister à des réunions du projet BOOSTER pour évaluer dans leurs discussions, la place du principe de garantie humaine et les représentations de ses enjeux. Nos résultats se basent donc uniquement sur la tenue d'entretiens semi-directifs. L'observation est pourtant le seul outil méthodologique qui permet « *d'accéder à certaines dimensions du social.* » (70)

Une autre limite tient à la construction de la grille d'entretien. Les premiers entretiens nous ont permis de constater que certaines questions obligeaient les personnes interrogées à se répéter. Ces premiers entretiens ont influencé la tenue des entretiens suivants. Le fait que l'enquête et l'analyse des entretiens soient à la charge de la même personne, associé au caractère flexible de la grille d'entretien (qui a donc évolué au cours du temps) peut être considéré comme source de biais.

Proposer une vignette clinique en fin d'entretien était une façon de compenser le manque d'observation. L'objectif était d'analyser les réactions des professionnels de santé dans une mise en situation fictive. Cet objectif n'a que partiellement été atteint. Pour certains, la vignette permettait de revenir sur des propos précédents échangés lors de l'entretien, de les renforcer, de faire émerger des concepts nouveaux. Pour d'autres, la vignette n'était pas pertinente, car le cas clinique n'était pas assez crédible.

La grande majorité des entretiens s'est déroulée par visioconférence en respect des mesures sanitaires. Il nous a semblé perdre en authenticité, en lien social. Il était plus difficile d'instaurer une confiance pour que la personne se sente à l'aise pour communiquer. La qualité des échanges a-t-elle été impactée ? Sans doute.

Reste désormais à s'interroger sur l'impact du point de vue juridique de l'enquêteur, complexifiant sans nul doute l'abord des professionnels et du sujet laissant parfois trop de

place au caractère médical et technique lors de l'entretien sans pouvoir complètement saisir les enjeux sous-jacents.

### **5.3.2 Points forts de l'étude**

Malgré ses limites, on peut reconnaître à cette étude un intérêt : celui de mettre en lumière les enjeux d'un principe complexe, dont le processus et les déterminants n'étaient pas évidents puisqu'exempts de tout cadre juridique prédéfini.

Nous pensons que la deuxième force de cette recherche est le solide arrière-plan théorique et conceptuel pour appuyer notre analyse. Notre question de départ était la suivante : quel est le lien entre le principe de garantie humaine du numérique en santé et le développement des applications de l'IA en santé ? Quelles interactions pouvons-nous établir entre ces deux concepts ? Comment les législateurs en sont-ils venus à proposer l'intégration de ce nouveau principe dans la révision de la loi de bioéthique ? Pourquoi dans la loi de bioéthique alors que l'IA n'a jamais fait partie des thèmes abordés par cette loi ? Le principe de garantie humaine répond-il aux enjeux posés par le développement de l'IA en santé ? Est-il le moyen le plus pertinent, le plus adapté pour répondre à ces enjeux ? Nous pensons que concevoir un principe juridique coupé de toute histoire n'est pas pertinent : nous avons alors mêlé les apports des disciplines juridiques, philosophiques, scientifiques et la science-fiction afin de déterminer les cadres de pensée à l'intérieur desquels il s'insère. Par conséquent, nous avons participé à plusieurs colloques, webinaires sur l'IA en santé et sur le principe de garantie humaine. Sans compter que nous avons eu l'opportunité d'échanger directement avec Monsieur David Gruson sur le principe de garantie humaine. Tout ceci a enrichi notre arrière-plan théorique.

Notre échantillonnage ne s'est pas limité aux participants du projet BOOSTER. Nous avons élargi notre panel de participants à des concepteurs d'outils d'IA et des professionnels de santé extérieurs au projet BOOSTER. Leurs visions ont permis de corroborer, de compléter et parfois d'opposer les différents points de vue. En outre, nous pensions au départ interroger uniquement des concepteurs d'outils d'IA (au sens de développeurs d'outils d'aide à la décision intégrant de l'IA). Bien que le milieu médical nous soit complètement extérieur,



nous avons fait l'effort de recruter des professionnels de santé concernés par le développement de ces outils. L'ajout d'une vignette clinique a permis dans certains cas d'échanger plus concrètement autour des enjeux de la responsabilité et du rôle de chacun (concepteur, machine, médecin par exemple). De nouveaux types de responsabilités que nous n'avions pas envisagés au départ sont venus enrichir les résultats, par exemple la responsabilité du testeur de l'outil d'IA. Un participant a même imaginé, à l'occasion de la vignette, un système entier de supervision humaine en aval de la conception de l'outil (« *Collège de médecins* », E7, 1.251).

Le caractère évolutif de la grille d'entretien nous a permis d'adapter les questions pour qu'elles soient plus pertinentes et nous rapprocher un peu plus de la question de recherche. Grâce à cette flexibilité, nous avons pu rebondir sur certains propos des participants à l'aide de questions qui n'étaient au départ pas prévues tout en respectant le temps qui nous était accordé par les participants. Par ailleurs, la transparence en début d'entretien sur notre formation juridique, et non médicale, a facilité les échanges, car les participants se sont adaptés avec bienveillance. Les discussions étaient enrichies par les parcours de chacun (médical, informatique, juridique, philosophique, économique, etc.). Un même enjeu pouvait ainsi être abordé sous plusieurs de ces angles. D'une certaine façon, cette extériorité au monde médical s'est révélée être un atout dans la mesure où nous ne portions aucun jugement sur leurs pratiques. Certes, nous avions quelques préjugés sur le monde médical et sur l'IA en général, mais les premiers échanges ont rapidement permis de les déconstruire. Par exemple, nous pensions qu'aucun concepteur n'allait envisager sa propre responsabilité en cas d'erreur médicale (hypothèse 2) et les résultats ont démontré l'inverse ! Les participants se sont éventuellement sentis plus à l'aise, nous l'espérons, car ils n'y avaient pas de rapport professionnel hiérarchique entre nous, pas de connaissances communes de collègues. Chacun venait d'une discipline différente et respectait cette différence.

## Conclusion

---

Le principe de garantie humaine est riche de complexité, parce qu'il questionne les applications nécessaires de l'intelligence artificielle en santé que personne ne remet en question, mais dont l'application apparaît compliquée.

Le point de départ de cette recherche était une double interrogation des plus concrète : d'une part, en quoi consiste le principe de garantie humaine du numérique en santé ? D'autre part, comment ce principe peut-il s'inscrire dans le développement des applications de l'IA ?

Le principe de garantie humaine est un début de réflexion éthique sur les enjeux du développement de l'IA en santé. L'essor de l'IA estompe progressivement la distinction entre l'Homme et la machine, remet en cause le rôle du médecin, sinon de la médecine, tandis que les données de santé participent, avec d'autres, au phénomène des mégadonnées (10). La régulation de l'IA en santé est « *un enjeu de souveraineté nationale dans un contexte de course technologique internationale* » (4).

Plus qu'un principe juridique en construction, la garantie humaine interroge sur le type de société dans lequel nous souhaitons évoluer. Il rappelle que la place de l'humain dans le soin n'est pas évidente et qu'il faut repenser le soin pour la préserver.

L'article 11 du projet de loi bioéthique est une illustration de la place du Droit dans la réflexion éthique : si le Droit s'abstient de tout réguler, il laisse une porte ouverte à des encadrements futurs. La réflexion éthique n'est pas la réflexion juridique.

Il ressort également de l'étude que la construction juridique telle que prévue par le législateur en 1978 à l'article 10 de la loi dite « Informatique et libertés » (50) ne se suffit pas à elle-même pour justifier l'interdiction d'un diagnostic ou d'une décision médicaux établis uniquement par un système d'IA sans intervention conjointe d'un médecin.

Au terme de cette étude, un certain nombre d'éléments semblent en défaut par rapport à la rédaction de l'article 11 du projet de loi relatif à la bioéthique. Le message clef au

législateur sur la base des résultats de recherche serait le suivant : le principe de garantie humaine du numérique en santé s'articule autour d'une supervision humaine, d'une relation humaine dans le soin et d'une responsabilité humaine.

1. **SUPERVISION HUMAINE** : Une équipe pluridisciplinaire composée des différentes parties prenantes (professionnels de santé, représentants de patients, politiques, philosophes, sociologues, etc.) doit mener une réflexion quant à la place à accorder à l'IA en santé. Une décision médicale ne peut pas dépendre du seul résultat brut issu d'un traitement automatisé de données massives. Le médecin doit pouvoir interpréter ce résultat à la lumière des autres éléments cliniques dont il est le seul à disposer. L'IA doit être considérée comme un outil complémentaire à l'activité du médecin sans jamais s'y substituer.
2. **RELATION HUMAINE** : Si l'IA peut se révéler être plus performante techniquement que le médecin, il ne faut jamais perdre de vue la dimension humaine qu'apporte le médecin au patient dans le soin. En effet, une machine est dépourvue des capacités d'empathie, de sensibilité et d'adaptation à des situations imprévues par l'algorithme. Les résultats techniques ne doivent pas prendre le pas sur la relation humaine. Prévoir un droit d'information au patient du recourt à un outil d'IA implique pour qu'il soit effectif, de lui réserver un droit d'opposition à cet usage. Or, il ne nous semble pas que le patient soit à même de décider s'il convient ou non de recourir à l'IA. En effet, cela relève bien plus de l'autonomie des professionnels de santé, laquelle repose sur une liberté de prescription et une liberté de jugement.
3. **RESPONSABILITÉ HUMAINE** : D'une part, la liberté de choix du professionnel de santé de recourir à de l'IA ne doit pas être absolue. Si le médecin décidait de ne pas recourir à l'IA et qu'un lien de causalité pouvait être établi entre ce choix et l'erreur médicale, alors il engagerait sa responsabilité. D'autre part, pour l'hypothèse où le médecin a recours à l'IA et qu'une erreur médicale due à un dysfonctionnement de la machine, dont le médecin ne pouvait avoir connaissance, est commise, il faut prévoir un nouveau régime de responsabilité spécifique à l'utilisation de l'IA en santé. Il ne nous semble pas éthiquement souhaitable d'attribuer une responsabilité juridique à l'IA. Cela dédouanerait tous les êtres humains impliqués dans le processus de création et d'utilisation de l'IA, d'une potentielle responsabilité.

Reste désormais à s'interroger sur l'essence de l'être humain pour comprendre ce qui justifie une garantie humaine face à des prouesses technologiques. Pourquoi l'IA, un outil et non pas un concurrent, ne peut-elle en aucun cas se substituer à un médecin ? Ces outils, ne relèvent-ils pas plus d'une Informatique Augmentée que d'une Intelligence Artificielle qui se prétendrait l'égale de l'Homme ? Quel est ce propre de l'être humain, si précieux, que la machine ne pourra jamais égaler ?

## Bibliographie

---

1. Comité consultatif national d'éthique pour les sciences de la vie et de la santé. Rapport de synthèse du Comité consultatif national d'éthique: opinions citoyennes (version intégrale) [Internet]. 2018 juin [cité 15 avr. 2021]. Disponible sur: [https://www.ccne-ethique.fr/sites/default/files/eg\\_ethique\\_rapportbd.pdf](https://www.ccne-ethique.fr/sites/default/files/eg_ethique_rapportbd.pdf)
2. Gruson D, Grass É. Quelle régulation positive éthique de l'intelligence artificielle en santé ? Trib Santé. 8 mai 2020;N° 63(1):25-33.
3. Gruson D. Le numérique et l'intelligence artificielle en santé : surveillance généralisée ou avancée majeure ? Trib Santé. 30 août 2019;N° 60(2):23-9.
4. Villani C. Donner un sens à l'intelligence artificielle : pour une stratégie nationale et européenne [Internet]. Cabinet du Premier ministre; 2018 mars [cité 15 avr. 2021]. Disponible sur: <https://www.vie-publique.fr/sites/default/files/rapport/pdf/184000159.pdf>
5. Comité consultatif national d'éthique. Avis 129 : Contribution du Comité Consultatif National d'Éthique à la révision de la loi de bioéthique 2018-2019 (version intégrale) [Internet]. Comité consultatif national d'éthique; 2018 sept [cité 15 avr. 2021]. Report No.: Avis 129. Disponible sur: [https://www.ccne-ethique.fr/sites/default/files/avis\\_129\\_vf.pdf](https://www.ccne-ethique.fr/sites/default/files/avis_129_vf.pdf)
6. Sénat. Projet de loi relatif à la bioéthique [Internet]. 53 févr. 3, 2021. Disponible sur: <https://www.senat.fr/petite-loi-ameli/2020-2021/281.html>
7. European Commission. Proposal for a regulation of the European Parliament and of the Council. Laying down harmonised rules on artificial intelligence and amending certain union legislative acts. [Internet]. avr 21, 2021. Disponible sur: [file:///Users/patrickallouche/Downloads/regulation\\_ai\\_875509BF-C386-0D30-2CB7E56A798BA4EA\\_75788.pdf](file:///Users/patrickallouche/Downloads/regulation_ai_875509BF-C386-0D30-2CB7E56A798BA4EA_75788.pdf)
8. Ethik-IA. La Garantie Humaine de l'IA en santé : cadre juridique de l'obligation de conformité et méthodologie-pilote dans le cadre du projet UFSBD/Ethik-IA.
9. Comité consultatif national d'éthique pour les sciences de la vie et de la santé. Résumé du rapport de synthèse du Comité consultatif national d'éthique: opinions citoyennes [Internet]. 2018 juin [cité 15 avr. 2021]. Disponible sur: <https://www.ccne-ethique.fr/sites/default/files/resume.pdf>
10. Conseil d'État, section du rapport et des études. Étude à la demande du Premier ministre - Révision à la loi de bioéthique : quelles options pour demain ? [Internet]. Conseil

d'État; 2018 juin [cité 15 avr. 2021]. Disponible sur: <https://www.conseil-etat.fr/ressources/etudes-publications/rapports-etudes/etudes/revision-de-la-loi-de-bioethique-quelles-options-pour-demain>

11. Jahr F. Aufsätze zur Bioethik 1924-1948. In: LIT Verlag [Internet]. Lit Verlag. Berlin; 2013 [cité 27 avr. 2021]. p. p°153. Disponible sur: <https://www.lit-verlag.de/isbn/978-3-643-12137-0>

12. Braibant G, France, Conseil d'État, Section du rapport et des études. De l'éthique au droit: sciences de la vie. Paris: La Documentation française; 1988.

13. Loi n° 88-1138 du 20 décembre 1988 relative à la protection des personnes qui se prêtent à des recherches biomédicales.

14. LOI n° 2011-814 du 7 juillet 2011 relative à la bioéthique (1). 2011-814 juill. 7, 2011.

15. Comité consultatif national d'éthique. Résumé de l'Avis 129 : Contribution du Comité Consultatif National d'Éthique à la révision de la loi de bioéthique [Internet]. Comité consultatif national d'éthique; 2018 sept [cité 15 avr. 2021]. Disponible sur: [https://www.ccne-ethique.fr/sites/default/files/resume\\_avis\\_129\\_vf\\_1010.pdf](https://www.ccne-ethique.fr/sites/default/files/resume_avis_129_vf_1010.pdf)

16. Décision n° 94-343-344 DC du 27 juillet 1994.

17. Conseil d'État. Avis sur un projet de loi relatif à la bioéthique [Internet]. NOR : SSAX1917211L juill. 18, 2019. Disponible sur: [https://www.legifrance.gouv.fr/contenu/Media/Files/autour-de-la-loi/legislatif-et-reglementaire/avis-du-ce/2019/avis\\_ce\\_ssax19172111\\_pjl\\_bioethique\\_cm\\_24.07.2019.pdf](https://www.legifrance.gouv.fr/contenu/Media/Files/autour-de-la-loi/legislatif-et-reglementaire/avis-du-ce/2019/avis_ce_ssax19172111_pjl_bioethique_cm_24.07.2019.pdf)

18. Loi n° 94-653 du 29 juillet 1994 relative au respect du corps humain.

19. Numérique. In: Trésor de la langue française informatisé [Internet]. CNRTL; 1994 [cité 11 mai 2021]. Disponible sur: <https://www.cnrtl.fr/definition/num%C3%A9rique>

20. OMS. Préambule à la Constitution de l'Organisation mondiale de la Santé, tel qu'adopté par la Conférence internationale sur la Santé, New York, 19 juin -22 juillet 1946; signé le 22 juillet 1946 par les représentants de 61 États. (Actes officiels de l'Organisation mondiale de la Santé, n°. 2, p. 100) et entré en vigueur le 7 avril 1948 [Internet]. Disponible sur: <https://www.who.int/fr/about/who-we-are/frequently-asked-questions>

21. Kempf A. L'informatisation de l'aide à la décision: la décision médicale est-elle indemne ? L'exemple d'un outil prédictif en cancérologie. Rev Fr Déthique Appliquée. 2016;1(1):59-70.

22. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nat Lond. 2017;542(7639):115-8.

23. HAS. Certification des logiciels d'aide à la prescription, une démarche primordiale pour l'amélioration des pratiques des médecins [Internet]. 2018 [cité 28 avr. 2021]. Disponible sur: [https://www.has-sante.fr/jcms/c\\_2864746/fr/certification-des-logiciels-d-aide-a-la-prescription-une-demarche-primordiale-pour-l-amelioration-des-pratiques-des-medecins](https://www.has-sante.fr/jcms/c_2864746/fr/certification-des-logiciels-d-aide-a-la-prescription-une-demarche-primordiale-pour-l-amelioration-des-pratiques-des-medecins)
24. Comité consultatif national d'éthique. Avis 109 : Communication d'informations scientifiques et médicales, et société : enjeux éthiques [Internet]. Comité consultatif national d'éthique; 2010 févr. [cité 15 avr. 2021]. Report No.: Avis 109. Disponible sur: <https://www.ccne-ethique.fr/fr/publications/communication-dinformatons-scientifiques-et-medicales-et-societe-enjeux-ethiques>
25. Ce qu'il faut savoir sur Cambridge Analytica, la société au cœur du scandale Facebook. Le Monde.fr [Internet]. 22 mars 2018 [cité 11 mai 2021]; Disponible sur: [https://www.lemonde.fr/pixels/article/2018/03/22/ce-qu-il-faut-savoir-sur-cambridge-analytica-la-societe-au-c-ur-du-scandale-facebook\\_5274804\\_4408996.html](https://www.lemonde.fr/pixels/article/2018/03/22/ce-qu-il-faut-savoir-sur-cambridge-analytica-la-societe-au-c-ur-du-scandale-facebook_5274804_4408996.html)
26. Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données) (Texte présentant de l'intérêt pour l'EEE) [Internet]. OJ L, 32016R0679 mai 4, 2016. Disponible sur: <http://data.europa.eu/eli/reg/2016/679/oj/fra>
27. CNIL. Qu'est-ce ce qu'une donnée de santé ? [Internet]. [cité 11 mai 2021]. Disponible sur: <https://www.cnil.fr/fr/quest-ce-ce-quune-donnee-de-sante>
28. Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.
29. Loi n° 2002-303 du 4 mars 2002 relative aux droits des malades et à la qualité du système de santé (1).
30. LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique (1). 2016-1321 oct. 7, 2016.
31. LOI n° 2018-493 du 20 juin 2018 relative à la protection des données personnelles (1). 2018-493 juin 20, 2018.
32. Article L1110-4 - Code de la santé publique [Internet]. Code de la santé publique. Disponible sur: [https://www.legifrance.gouv.fr/codes/article\\_lc/LEGIARTI000036515027/](https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000036515027/)
33. Conseil national de l'ordre des médecins. Médecins et patients dans le monde des data, des algorithmes et de l'intelligence artificielle. [Internet]. 2018 [cité 28 avr. 2021]. Disponible sur: <https://www.conseil-national.medecin.fr/sites/default/files/external->

package/edition/od6gnt/cnomdata\_algorithmes\_ia\_0.pdf

34. Lacroux A, Martin-Lacroux C. L'Intelligence artificielle au service de la lutte contre les discriminations dans le recrutement : nouvelles promesses et nouveaux risques. *Manag Avenir*. 21 avr. 2021;N° 122(2):121-42.
35. FLORIAN DEBES. L'AP-HP victime d'une cyberattaque. *Échos* (Paris, France). 2020;
36. Leriche R. *La philosophie de la chirurgie*. Paris, France: Flammarion; 1951. 251 p.
37. Demiaux V, Si Abdallah Y. Comment permettre à l'Homme de garder la main ? Les enjeux éthiques des algorithmes et de l'intelligence artificielle [Internet]. CNIL; 2017 déc. [cité 15 avr. 2021]. Disponible sur: [https://www.cnil.fr/sites/default/files/atoms/files/cnil\\_rapport\\_garder\\_la\\_main\\_web.pdf](https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_garder_la_main_web.pdf)
38. Assemblée Nationale, Breton X, Touraine J-L. Rapport d'information fait au nom de la mission d'information sur la révision de la loi relative à la bioéthique [Internet]. Assemblée nationale; 2019 janv. [cité 15 avr. 2021]. Report No.: N°1572. Disponible sur: [https://www.assemblee-nationale.fr/dyn/15/rapports/bioethique/115b1572\\_rapport-information](https://www.assemblee-nationale.fr/dyn/15/rapports/bioethique/115b1572_rapport-information)
39. Lai M-C, Brian M, Mamzer M-F. Perceptions of artificial intelligence in healthcare: Findings from a qualitative survey study among actors in France. *J Transl Med*. 2020;18(1):14-14.
40. Collectif C. Éthique de la recherche en apprentissage machine [Internet] [report]. CERNA; ALLISTENE; 2017 [cité 30 avr. 2021]. p. 51. Disponible sur: <https://hal.inria.fr/hal-01643281>
41. Ethik-IA. La Garantie Humaine dans le projet de règlement sur l'IA de la Commission européenne ! 2021.
42. CARE INSIGHT. Conformité en garantie humaine de l'IA en santé : Comment s'y préparer ? In 2021. Disponible sur: [https://careinsight.clickmeeting.com/conformite-en-garantie-humaine-de-l-ia-en-sante-comment-s-y-preparer-?r=AGL5AwV0ZQq8sSqbrKMIVR55rJWbpUIIsUk3nUy2pz55rGVkDTkhqJWvYaAysUjmZGHjAGtkZj\\_\\_\\_\\_\\_](https://careinsight.clickmeeting.com/conformite-en-garantie-humaine-de-l-ia-en-sante-comment-s-y-preparer-?r=AGL5AwV0ZQq8sSqbrKMIVR55rJWbpUIIsUk3nUy2pz55rGVkDTkhqJWvYaAysUjmZGHjAGtkZj_____)
43. Gruson D. S.A.R.R.A. une intelligence artificielle [Internet]. BETA PUBLISHER. 2018 [cité 30 avr. 2021]. 320 p. Disponible sur: <https://www.babelio.com/livres/Gruson-SARRA--Une-intelligence-artificielle/1049983>
44. Gruson D. Échanges avec Monsieur David Gruson sur le principe de garantie humaine. 2021.



45. CNIL. Délibération de la Commission Nationale de l'Informatique et des Libertés [Internet]. 2019-097 juill. 11, 2019. Disponible sur: <https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000038848207/>
46. Assemblée nationale. Étude d'impact : Projet de loi relatif à la bioéthique [Internet]. NOR : SSAX1917211L/Bleue-1 juill. 23, 2019. Disponible sur: [https://www.legifrance.gouv.fr/contenu/Media/Files/autour-de-la-loi/legislatif-et-reglementaire/etudes-d-impact-des-lois/ei\\_art\\_39\\_2019/ei\\_ssax19172111\\_bioethique\\_cm\\_24.07.2019.pdf](https://www.legifrance.gouv.fr/contenu/Media/Files/autour-de-la-loi/legislatif-et-reglementaire/etudes-d-impact-des-lois/ei_art_39_2019/ei_ssax19172111_bioethique_cm_24.07.2019.pdf)
47. Comité Consultatif National d'Éthique. Avis 130 : Données massives et santé. Une nouvelle approche des enjeux éthiques. [Internet]. 2019 mai [cité 6 juin 2021]. Report No.: 130. Disponible sur: [https://www.ccne-ethique.fr/sites/default/files/avis\\_130.pdf](https://www.ccne-ethique.fr/sites/default/files/avis_130.pdf)
48. Groupe de travail CCNE/CERNA. Numérique et santé : quels enjeux éthiques pour quelles régulations ? [Internet]. 2018 juill. [cité 30 avr. 2021] p. 100. Disponible sur: [https://www.allistene.fr/files/2018/11/rapport\\_numerique\\_et\\_sante\\_19112018.pdf](https://www.allistene.fr/files/2018/11/rapport_numerique_et_sante_19112018.pdf)
49. Article L4161-1 - Code de la santé publique - Légifrance [Internet]. Code de la santé publique. Disponible sur: [https://www.legifrance.gouv.fr/codes/article\\_lc/LEGIARTI000038886735/](https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000038886735/)
50. Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.
51. Article L1111-2 - Code de la santé publique - Légifrance [Internet]. Code de la santé publique. Disponible sur: [https://www.legifrance.gouv.fr/codes/article\\_lc/LEGIARTI000041721051/](https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000041721051/)
52. Article L1111-4 - Code de la santé publique - Légifrance [Internet]. Disponible sur: [https://www.legifrance.gouv.fr/codes/article\\_lc/LEGIARTI000041721056/](https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000041721056/)
53. Article 45 - Constitution du 4 octobre 1958 [Internet]. Disponible sur: [https://www.legifrance.gouv.fr/loda/article\\_lc/LEGIARTI000019241040/#:~:text=Tout%20projet%20ou%20proposition%20de,adoption%20d'un%20texte%20identique.&text=Le%20texte%20%C3%A9labor%C3%A9%20par%20la,pour%20approbation%20aux%20deux%20Assembl%C3%A9es.](https://www.legifrance.gouv.fr/loda/article_lc/LEGIARTI000019241040/#:~:text=Tout%20projet%20ou%20proposition%20de,adoption%20d'un%20texte%20identique.&text=Le%20texte%20%C3%A9labor%C3%A9%20par%20la,pour%20approbation%20aux%20deux%20Assembl%C3%A9es.)
54. Sénat. Amendement d'Alain Milon adopté en 2de lecture au Sénat après favorable du gouvernement [Internet]. Disponible sur: [https://www.senat.fr/compte-rendu-commissions/20210201/cs\\_bioethique.html](https://www.senat.fr/compte-rendu-commissions/20210201/cs_bioethique.html)
55. Haute Autorité de santé (HAS). Un nouvel outil pour l'évaluation des dispositifs médicaux embarquant de l'intelligence artificielle [Internet]. Haute Autorité de Santé. 2021 [cité 2 juin 2021]. Disponible sur: [https://www.has-sante.fr/jcms/p\\_3212876/fr/un-nouvel-](https://www.has-sante.fr/jcms/p_3212876/fr/un-nouvel-)

outil-pour-l-evaluation-des-dispositifs-medicaux-embarquant-de-l-intelligence-artificielle

56. Commission européenne. Livre Blanc : Intelligence artificielle. Une approche européenne axée sur l'excellence et la confiance. [Internet]. Bruxelles; 2020 [cité 2 juin 2021]. Disponible sur: [https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020\\_fr.pdf](https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_fr.pdf)
57. Beauchamp Tom L., Fissbach Martine. Les principes de l'éthique biomédicale. Paris: Les Belles Lettres DL 2008; (Médecine & sciences humaines).
58. Emanuel EJ, Emanuel LL. Four Models of the Physician-Patient Relationship. JAMA. 22 avr. 1992;267(16):2221-6.
59. Paillé P, Mucchielli A. L'analyse qualitative en sciences humaines et sociales. 4e éd. Armand Colin; 2016. (U).
60. Baribeau C, Royer C. L'entretien individuel en recherche qualitative: usages et modes de présentation dans la Revue des sciences de l'éducation. Rev Sci L'éducation. 2012;38(1):23-45.
61. Bardin Laurence. L'analyse de contenu / Laurence Bardin. Paris: PUF; 2017. (Quadrige).
62. HAS. Méthode de DPC - Vignettes cliniques [Internet]. Haute Autorité de Santé. [cité 11 mai 2021]. Disponible sur: [https://www.has-sante.fr/jcms/p\\_3201597/fr/methode-de-dpc-vignettes-cliniques](https://www.has-sante.fr/jcms/p_3201597/fr/methode-de-dpc-vignettes-cliniques)
63. Lebeau Jean-Pierre. Initiation à la recherche qualitative en santé: Le guide pour réussir sa thèse ou son mémoire / Sous la direction de Jean-Pierre Lebeau, Isabelle Aubin-Auger, Jean-Sébastien Cadwallader... [et al.]; Préface Paul Frappé. Mayenne: Global media santé; 2021. 192 p.
64. Antoine P, Smith JA. Saisir l'expérience: présentation de l'analyse phénoménologique interprétative comme méthodologie qualitative en psychologie. Psychol Fr. 2017;62(4):373-85.
65. dit J. La Télémedecine - TIC santé et Téléconsultation - définition [Internet]. France Assos Santé. [cité 3 juin 2021]. Disponible sur: <https://www.france-assos-sante.org/66-millions-dimpatients/la-qualite-de-vos-soins/la-telemedecine/>
66. Jaunait Alexandre, Sadoun Marc. Comment pense l'institution médicale ? : une analyse des codes français de déontologie médicale. [Paris]: Dalloz; 2005.
67. Parlement européen. Résolution du Parlement européen contenant des recommandations à la Commission concernant les règles de droit civil sur la robotique [Internet]. Févr. 16, 2017. Disponible sur:

[https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051\\_FR.html](https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_FR.html)

68. Toll E. The Cost of Technology. JAMA. 20 juin 2012;307(23):2497-8.
69. Illich Ivan. La convivialité. Paris: Éditions du Seuil; 1973.
70. Fabre D. L'ethnologue et ses sources. Terrain Paris 1983. 1986;(7):3-12.

## Table des figures

---

Figure 1 : Sources des données de santé.....	11
Figure 2 : Enjeux éthiques du numérique et santé .....	15
Figure 3 : Les principes de régulation de l'IA .....	17
Figure 4 : Chronologie du principe de garantie humaine .....	21
Figure 5 : Les étapes de la discussion du projet de loi relatif à la bioéthique .....	30
Figure 6 : Introduction du principe de garantie humaine à l'international .....	37
Figure 7 : Un dessin d'enfant représentant la consultation avec le médecin à l'ère des ordinateurs portables dans la salle d'examen .....	76

## Table des tableaux

---

Tableau 1 : Régimes juridiques applicables aux données de santé.....	12
Tableau 2 : Connaissance du principe de garantie humaine .....	51
Tableau 3 : Types de responsabilité envisagés par les participants .....	52
Tableau 4 : Responsabilité du concepteur en cas d'erreur médicale .....	68
Tableau 5 : Responsabilité du médecin en cas d'erreur médicale .....	70

## **Annexes**

---

**Annexe 1** : Récapitulatif historique en matière de législation dans le domaine de la bioéthique, de la loi de 1988 à sa dernière révision en 2011.

**Annexe 2** : Lettre de mission du Premier ministre à Monsieur le vice-président du Conseil d'État, le 06/12/2017.

**Annexe 3** : Grille d'entretien et vignette clinique.

**Annexe 4** : E-mail type de recrutement.

**Annexe 5** : Lettre d'information à la recherche.

## Annexe 1 : Récapitulatif historique en matière de législation dans le domaine de la bioéthique, de la loi de 1988 à sa dernière révision en 2011

RAPPORT DE SYNTHÈSE DU COMITÉ CONSULTATIF NATIONAL D'ÉTHIQUE

Figure 1: Récapitulatif historique

1988	Loi n° 88-1138	Thèmes abordés: Protection de la personne dans le cadre des essais cliniques (recherche biomédicale). Évolutions institutionnelles: Création des CCPPRB (Comité consultatif de protection des personnes dans la recherche biomédicale), aujourd'hui remplacés par les CPP (Comité de protection des personnes).	
1988	Rapport BRAIBANT: Sciences de la vie: de l'éthique au droit (étude du Conseil d'État)		Textes fondateurs de l'éthique dans la recherche biomédicale, introduisant notamment le concept de « bilan risques/ bénéfices ».
1991	Rapport LENOIR: Aux frontières de la vie: une éthique biomédicale à la française (rapport au Premier ministre)		
1993	Rapport MATTEI: Rapport sur l'éthique biomédicale (rapport au Premier ministre)		
	Loi n° 94-548	Thèmes abordés: Traitement des données de santé.	
1994	Loi n° 94-653	Thèmes abordés: Inviolabilité du corps humain, non-patrimonialité du corps humain, obligation du consentement.	
	Loi n° 94-654	Thèmes abordés: Dons de produits humains, AMP, diagnostic prénatal, principe de gratuité, d'anonymat, de respect des règles de sécurité sanitaire.	
2004	Loi n°2004-800	Thèmes abordés: Actualisation de la loi de 1994, thérapie cellulaire, dons d'organes, recherche en génétique, diagnostic préimplantatoire, recherche sur les embryons. Évolutions institutionnelles: Création de l'Agence de la biomédecine.	
2009		Consultations institutionnelles: Conseil d'État, CCNE, Agence de la biomédecine, OPECST. Consultations citoyennes: Organisation d'États généraux de la bioéthique à l'origine d'un certain nombre de débats avec le public et de jurys citoyens: Rennes (procréation), Strasbourg (don d'organes), Marseille (cellules souches).	
2011	Loi n° 2011-814	Thèmes abordés: AMP, don de gamètes, diagnostic anténatal, IMG, recherche sur l'embryon humain et les cellules souches embryonnaires, génétique, neurosciences, don d'organes. Évolutions institutionnelles: Élargissement des missions du CCNE.	
2013	Loi n° 2013-715	Thème abordé: recherche sur l'embryon.	
2018		Le législateur ayant inscrit au sein de la loi de 2011 la nécessité de procéder à son examen d'ensemble par le Parlement dans un délai maximum de sept ans après son entrée en vigueur, le Comité consultatif national d'éthique (CCNE) a officiellement ouvert le 18 janvier 2018 les États généraux de la bioéthique, préalable à la révision de la loi.	

## Annexe 2 : Lettre de mission du Premier ministre à Monsieur le vice-président du Conseil d'État, le 06/12/2017.

### 1.1. Lettre de mission du Premier ministre

---



*Le Premier Ministre*  
2200 / 17 / SG

Paris, le 06 DEC 2017

à

Monsieur le vice-président  
du Conseil d'Etat

Objet : Cadrage juridique préalable au réexamen de la loi relative à la bioéthique

La loi n° 2011-814 du 7 juillet 2011 relative à la bioéthique fera l'objet d'un nouvel examen d'ensemble par le Parlement, comme prévu à son article 47, « dans un délai maximal de sept ans après son entrée en vigueur ». Depuis la loi n° 2004-800 du 6 août 2004 relative à la bioéthique, le législateur souhaite en effet prendre en compte les avancées scientifiques et l'évolution de plus en plus rapide des technologies biomédicales, dans un contexte de mondialisation des enjeux de santé et de recherche, ainsi que les nouvelles demandes de la société induites par ces progrès et innovations médicales et scientifiques.

Les innovations technologiques et les découvertes scientifiques, par leur impact sur le système de santé, sur l'organisation sociale et les droits de la personne humaine, posent plus que jamais des questions éthiques et nous invitent à une réflexion collective. Le législateur a ainsi prévu que « tout projet de réforme sur les problèmes éthiques et les questions de société soulevés par les progrès de la connaissance dans les domaines de biologie, de la médecine et de la santé doit être précédé d'un débat public sous formes d'Etat généraux » (article 46), organisés à l'initiative du Comité consultatif national d'éthique pour les sciences de la vie et de la santé (CCNE). Ces Etats généraux se dérouleront fin 2017-début 2018.

Page 215





Cette démarche éthique et de débats publics doit être complétée par une réflexion approfondie sur les principes juridiques fondateurs établis en 1994, date de la première législation sur la bioéthique. Ce premier corpus juridique très complet, issu de trois lois complémentaires, a posé un cadre à la fois dans le code de la santé publique et dans le code civil. Il comporte également un volet relatif à la protection des données dans le champ de la recherche. Ces principes ont été confirmés lors des deux premières révisions de 2004 et 2011.

Le Conseil d'État a apporté au Gouvernement un précieux concours lors de l'élaboration des lois de 1994, 2004 et 2011 et je souhaite qu'il puisse contribuer à la réflexion préalable au prochain réexamen de la loi relative à la bioéthique, par des éléments de cadrage juridique.

La contribution du Conseil d'État pourrait porter sur l'identification et la formulation des problématiques juridiques qui lui apparaîtraient importantes, notamment celles pouvant conduire à une évolution de la législation, sur l'ensemble du champ relevant de l'éthique, y compris au-delà des thèmes et sujets d'ores et déjà pressentis. À ce stade, six thématiques ont été identifiées en lien avec le Comité consultatif national d'éthique, au sein desquelles certains sujets plus précis se distinguent dans les champs de réflexion suivants :

- La procréation : notamment les questions de l'accès à l'assistance médicale à la procréation par des couples de femmes et des femmes seules ; l'autoconservation ovocytaire chez les femmes et la conservation de cellules du sang de cordon pour un éventuel usage ultérieur notamment dans le cadre familial ; l'« accès aux origines » pour les enfants nés par assistance médicale à la procréation avec don de gamètes ; l'insémination artificielle et le transfert embryonnaire après la mort du donneur de gamètes ; l'accès à la gestation pour autrui ; plus généralement, les enjeux posés par les différentes techniques de procréation en droit de la filiation.
- Les conditions du don d'organes, de tissus et de cellules (dont les gamètes) et du don du sang : notamment l'anonymat, le principe de gratuité et la question de l'équilibre entre neutralité financière et prise en charge des contraintes liées au don ; l'extension du champ des donneurs vivants pour mieux répondre aux besoins d'organes.
- Les principes bioéthiques confrontés aux évolutions de la génomique, tels : la recherche sur l'embryon et les cellules souches embryonnaires humaines et les techniques d'édition du génome ; les examens des caractéristiques génétiques à haut débit, pouvant donner lieu à des découvertes incidentes ; la médecine génomique et la question d'un consentement libre et éclairé dans un contexte d'effacement des frontières entre prise en charge clinique et recherche.
- Les enjeux éthiques et juridiques émergents liés au développement de l'intelligence artificielle appliqué au domaine médical : constitution et utilisation du « *big data* », encadrement des finalités des biotechnologies, du numérique et des neurosciences.



- Les droits des malades et des personnes en fin de vie : l'adéquation de la loi actuelle avec les attentes sociétales, notamment la demande d'une aide médicale à mourir.

- La situation des enfants intersexes sur le plan biomédical et juridique : la question de la mention du sexe à l'état civil et celle des traitements médicaux d'assignation sexuelle.

Ce recensement, non exhaustif, aura également vocation à être enrichi par le Comité consultatif national d'éthique d'ici la fin de cette année.

J'attacherais du prix à ce que les réflexions et recommandations du Conseil d'État, assorties le cas échéant de propositions de modification législative, me soient remises avant la fin du premier semestre 2018. En effet, le dépôt d'un projet de loi par le Gouvernement, sur lequel le Conseil d'État sera saisi, est envisagé pour le dernier trimestre 2018.

  
Edouard PHILIPPE



## Annexe 3 : Grille d'entretien et vignette clinique

### Grille d'entretien

#### Le principe de garantie humaine du numérique en santé

Julie Allouche – Master 2, Université de Paris

#### Question de recherche :

Comment le principe de garantie humaine du numérique en santé peut-il s'inscrire dans le développement des applications de l'intelligence artificielle en santé ?

#### I. Présentation

- **Modalités de présentation :** Je m'appelle Julie Allouche, je viens d'obtenir un double diplôme en droit franco-espagnol de l'Université Paris I Panthéon-Sorbonne. Cette année, je suis étudiante en Master 2 d'éthique médicale à la faculté de médecine de Paris au sein du Laboratoire d'Éthique Médicale et de l'Équipe recherche « ETREs », dirigé par Madame le Professeur Marie-France MAMZER.
- **Acceptation de la présence du chercheur :** Tout d'abord, je tiens à vous remercier d'avoir accepté de participer à cet entretien dans le cadre de cette recherche.
- **Objectif de cet entretien :** L'objectif de mon travail est de réfléchir d'un point de vue éthique à un principe juridique émergent : le principe de garantie humaine. On pourra y revenir. Dans un premier temps, j'aimerais échanger avec vous sur les applications de l'IA en santé.
- **Acceptation de l'enregistrement :** Accepteriez-vous d'être enregistré(e) ? Les enregistrements audios se feront via le logiciel Zoom ou bien via l'enregistreur vocal de l'ordinateur personnel de Julie Allouche. Les données seront extraites des enregistrements et seront stockées sur un ordinateur individuel sécurisé par un mot de passe. Elles seront aussi sauvegardées sur un disque dur dans un tiroir fermé à clef. Les données seront pseudonymisées et retranscrites par écrit. Les enregistrements seront alors détruits.
- **Modalités de l'entretien :** Sachez qu'aucun jugement ne sera fait durant notre échange. L'entretien a pour but que vous partagiez vos expériences personnelles et professionnelles. Possibilité de faire des pauses. Possibilité d'arrêter l'enregistrement dès que souhaité.

## II. Dispositif d'entretien

- 1) Quelle est votre expérience des outils d'aide à la décision utilisant de l'IA ?
- 2) Quels avantages attendez-vous de ces outils ? Y voyez-vous des limites ?
- 3) Quels sont les objectifs poursuivis par la création de ces outils ?
- 4) Quelles questions posent ces outils ?
- 5) Connaissez-vous le principe de garantie humaine du numérique en santé ? [Si la personne ne connaît pas le principe :] Comment le comprendriez-vous intuitivement par rapport à vos connaissances de l'IA en santé ? [Si la personne me demande une définition du principe : « Parler d'une supervision humaine qui viendrait coiffer ces algorithmes. » Moi c'est ce que j'en ai compris, mais peut-être que j'ai mal compris, et vous qu'est-ce que vous en pensez ?]
- 6) La garantie humaine vous paraît-elle répondre à vos interrogations sur les applications de l'IA en santé ?
- 7) L'envisagez-vous comme : une source de limitation à la conception de ces outils ? Comme un cadre de protection de création et de production ?
- 8) Ce principe de garantie humaine vous paraît-il pertinent ? Quel système idéal imagineriez-vous ? Comment en pratique cela pourrait-il prendre forme ?

### **VIGNETTE clinique :**

En 2035, Monsieur T présente des signes compatibles avec un accident vasculaire cérébral (AVC). Son mari appelle le SAMU. Le médecin régulateur du SAMU consulte son dossier médical et retrouve les antécédents suivants : « patient atteint d'athérosclérose sévère présentant toutes les complications de la maladie : infarctus du myocarde avec portage

coronaire il y a 6 ans, artériopathie obstructive des membres inférieurs (AOMI) depuis 2 ans, 2 AVC ischémiques par sténose des carotides avec thrombus ayant nécessité une revascularisation ».

Une équipe SAMU est dépêchée sur place. L'examen clinique, réalisé par le médecin du SAMU, est en faveur d'un AVC sylvien droit. Le médecin renseigne le logiciel d'Intelligence artificielle de la symptomatologie. Aux vues des données cliniques, des antécédents et de l'épidémiologie de l'AVC (80% des cas sont ischémiques), le logiciel propose la mise en place d'un traitement par thrombolyse intraveineuse car l'hypothèse d'un AVC ischémique est retenue par l'algorithme. Il détermine aussi l'hôpital dans lequel Monsieur T doit être envoyé (en fonction de la distance, des places disponibles, du plateau technique et de l'expertise de chaque hôpital). Le médecin de l'ambulance procède à la thrombolyse.

20 minutes après la thrombolyse, Monsieur T décède dans l'ambulance. L'enquête de recherche des causes de la mort et les conclusions des experts conduisent à formuler une erreur de diagnostic initiale : le traitement était dangereux et a pu accélérer la mort. En effet, Monsieur T est décédé d'un AVC hémorragique sévère. La famille de Monsieur T souhaite porter plainte pour erreur médicale.

*Qu'en pensez-vous ?*

### III. CONCLUSION

- Proposer à la personne interrogée d'ajouter ou de revenir sur une information.
- Proposer exemplaire dactylographié anonymisé de cette observation.
- Proposer de partager les résultats de l'étude.
- Rappeler ma disponibilité et mes coordonnées + Possibilité de se retirer de l'étude à tout moment.
- Remercier !

## **Annexe 4 : E-mail type de recrutement**

Objet : Entretien intelligence artificielle

Bonjour,

Je m'appelle Julie Allouche et je m'intéresse, dans le cadre d'un Master 2 de recherche (au sein du laboratoire d'éthique médicale de l'Université de Paris) d'un point de vue éthique à un principe juridique émergent, le principe de garantie humaine du numérique en santé.

Ma grille d'entretien a été validée par mes directeurs de mémoire, le Professeur Marie-France Mamzer et Monsieur Dupont, et a déjà été testée lors d'entretiens préliminaires.

J'aimerais réaliser un entretien avec vous sur ce sujet. Je suis particulièrement intéressée par votre point de vue sur le développement d'algorithmes d'intelligence artificielle pour l'aide à la décision.

Vous pouvez me contacter par retour de mail pour que nous puissions convenir d'une date d'entretien par visioconférence.

Je suis par ailleurs disponible pour toute question complémentaire et joins la lettre d'information concernant cette recherche.

En espérant vivement votre collaboration,  
Cordialement,

Julie Allouche  
Master 2 d'éthique médicale de l'Université de Paris

## Annexe 5 : Lettre d'information à la recherche

### INFORMATION AUX SUJETS CONCERNANT LE PROTOCOLE DE RECHERCHE :

**Comment le principe de « garantie humaine » du numérique en santé peut-il s'inscrire dans le développement des applications de l'Intelligence artificielle ?**

Madame, Monsieur,

Cette recherche est menée par Julie ALLOUCHE et codirigée par le Pr. Marie-France MAMZER et Monsieur Jean-Claude K. DUPONT dans le cadre d'un Master 2 en éthique médicale au sein de l'équipe ETREs (Éthique, Recherche, Translations), Université de Paris, 15 rue de l'Ecole de médecine – 75006 PARIS, qui est responsable du traitement.

Les informations recueillies vous concernant vont faire l'objet d'un traitement dans le cadre du projet de recherche précité.

Votre participation à cette recherche est entièrement libre et volontaire, vous pouvez retirer votre consentement à tout moment en contactant Julie Allouche au [donnée personnelle retirée] ou par mail à [donnée personnelle retirée]. En cas de retrait de consentement, l'ensemble de vos données collectées dans le cadre de cette étude seront supprimées.

#### **Objectif de la recherche :**

Cette recherche a pour objectif de mieux comprendre le principe de garantie humaine, ses interactions avec l'intelligence artificielle (IA) en santé et de recueillir une vision des concepteurs de ces outils sur ce principe : leurs représentations sur les questionnements des législateurs autour des enjeux d'IA, ainsi que leur avis sur deux thématiques liées à leur potentielle responsabilité engagée en tant que concepteurs d'outils d'IA.

#### **Méthodologie de la recherche :**

Pour mener cette recherche, je souhaiterais, avec votre accord, réaliser un entretien avec vous, selon vos disponibilités durant la période qui s'étend de mi-Février à Avril 2021.

Cet entretien se déroulera si possible dans un environnement calme pour favoriser l'échange, peut-être en distanciel compte tenu des circonstances. Durant l'entretien, je vous poserai des questions en lien avec votre expérience en tant que chercheur/concepteur d'outils d'IA autour du concept de garantie humaine.

De plus, après avoir recueilli votre consentement libre et éclairé, je souhaiterai enregistrer notre entretien grâce à un dictaphone dédié à cette recherche. Les données seront stockées sur un ordinateur individuel sécurisé par un mot de passe.

A la suite de notre entretien, vos propos seront retranscrits fidèlement à l'écrit, en respectant votre anonymat et j'analyserai ensuite le contenu du récit obtenu. L'enregistrement de notre entretien sera détruit à la suite de cette retranscription.

#### **Protection des données à caractère personnel :**

Le traitement de vos données à caractère personnel se fera conformément au RÈGLEMENT (UE) 2016/679 DU PARLEMENT EUROPÉEN ET DU CONSEIL du 27 avril 2016 relatif à la protection

des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données.

Seules les données strictement nécessaires à la réalisation de notre projet de recherche seront collectées et traitées :

**Données sensibles :**

- Enregistrement de la voix pour la retranscription
- Les données de l'entretien qui portent sur le sujet de recherche

La base légale du traitement de vos données personnelles repose sur l'exécution d'une mission d'intérêt public dont est investi le responsable du traitement. Votre participation à la recherche est entièrement libre et volontaire. Il est rappelé que le participant est libre de retirer ou cesser sa participation à ce projet à tout moment. Ce retrait n'aura aucune conséquence.

Toutes les données personnelles collectées sont conservées en France.

**Les destinataires des données :**

Les données personnelles recueillies ne seront consultées que par Julie ALLOUCHE (ayant un master de droit franco-espagnol de l'Université Paris I Panthéon-Sorbonne et étudiante en Master 2 d'éthique médicale et de bioéthique), Monsieur Jean-Claude K. Dupont (directeur du travail, directeur adjoint de la chaire d'économie de la santé Hospinomics) et Madame Marie-France Mamzer (directrice du travail, responsable de l'équipe de recherche Université de Paris - UMRS 1138 Équipe ETREs – EA 4569).

Par ailleurs, il est possible que ce travail de recherche, fasse l'objet, dans un second temps, d'une publication sous forme d'article dans une revue internationale pour améliorer la connaissance scientifique sur le sujet étudié. Si tel est le cas, les résultats de la recherche seront diffusés de façon anonyme dans des colloques professionnels et scientifiques, dans des rapports destinés aux autorités, dans des revues professionnelles et académiques et dans des médias destinés au grand public.

**Durée de conservation des données :**

Vos données personnelles sont conservées en base active jusqu'à deux ans après la dernière publication des résultats, ou en cas d'absence de publication, jusqu'à la signature du rapport final de la recherche. Elles font ensuite l'objet d'un archivage sur support papier ou informatique pour une durée de deux ans après la fin de la recherche.

L'enregistrement vocal des entretiens sera supprimé une fois la retranscription sur un logiciel de traitement de texte réalisée.

Mesures de sécurité techniques et organisationnelles :

Afin de garantir la confidentialité de vos données et éviter leur divulgation, le responsable de la recherche prend toutes les précautions utiles pour préserver la sécurité des données traitées, en particulier leur confidentialité, leur intégrité et leur disponibilité. Pour ce faire, les dispositifs suivants ont été mis en place :

- Seules les personnes collaborant à la recherche, désignées par le responsable de la recherche ou par un représentant des autorités administratives compétentes sont autorisés à accéder aux données.



- Les mesures de sécurité, tant physique que logique, suivantes sont prises : sauvegarde sur un disque dur dans un tiroir fermé à clé, présence d'un logiciel antivirus.

La retranscription se fera sur un ordinateur protégé par un mot de passe et sera placée dans un dossier chiffré par un logiciel approuvé par l'Agence nationale de la sécurité des systèmes d'informations conformément aux bonnes pratiques recommandées par la CNIL.

**Vos droits :**

Conformément aux dispositions du Règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données et de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés et ses décrets d'application, vous pouvez accéder et obtenir copie des données vous concernant, vous opposer au traitement de ces données, les faire rectifier ou les faire effacer. Vous disposez également d'un droit à la limitation du traitement de vos données.

En cas d'exercice de votre droit d'opposition au traitement, vous pourrez demander l'effacement des données vous concernant déjà collectées. Cependant, vous êtes informé que certaines données préalablement collectées ne pourront être effacées, si cette suppression est susceptible de rendre impossible ou de compromettre gravement la réalisation des objectifs de la recherche.

Vous pouvez exercer ces droits ou poser des questions au sujet de cette recherche auprès du responsable de la recherche en vous adressant à Julie Allouche, à l'adresse suivante : [donnée personnelle retirée]. Une réponse vous sera apportée dans les plus brefs délais avec un délai maximal d'un mois à compter de la réception de la demande.

Vous pouvez contacter également le Délégué à la Protection des Données de l'Université de Paris à l'adresse postale suivante : DPD, 18 rue de l'Ecole de Médecine 75006 PARIS – ou à [donnée personnelle retirée].

Après nous avoir contactés, si vous estimez que vos droits Informatique et Libertés ne sont pas respectés, vous avez la possibilité d'introduire une réclamation en ligne auprès de la CNIL ou par courrier postal. CNIL, 3 Place de Fontenoy, TSA 80715 – 75334 Paris Cedex 07 (<https://www.cnil.fr>).

Si vous le souhaitez, vous pouvez demander des informations sur les résultats globaux de la recherche en contactant [donnée personnelle retirée].

Fait à Paris le .....

Signature de la personne sollicitée


## Autorisations de diffusion des œuvres utilisées

---

**Autorisation 1 :** Les étapes de la discussion du projet de loi relatif à la bioéthique, [www.senat.fr](http://www.senat.fr), 07/06/2021.

**Autorisation 2 :** Introduction du principe de Garantie Humaine à l'internationale, Health & Tech Intelligence – Care Insight, 01/06/2021.

**Autorisation 1** : Les étapes de la discussion du projet de loi relatif à la bioéthique,  
[www.senat.fr](http://www.senat.fr), 07/06/2021.

**De:** Cellule internet du Sénat cellule-internet@senat.fr   
**Objet:** Re: Mémoire de recherche - schéma navette parlementaire loi de bioéthique  
**Date:** 3 mai 2021 à 16:25  
**À:** Allouche Julie julieall21@yahoo.fr  
**Cc:** cellule-internet@senat.fr

---

Madame,

Les documents parlementaires sont régis par le principe de libre reproduction des actes officiels : ceux-ci ne sont couverts par aucun droit d'auteur (article L. 122-5 du code de la propriété intellectuelle).

La reproduction sous forme papier ou électronique est libre, à condition d'en mentionner le crédit, la citation expresse du site [www.senat.fr](http://www.senat.fr) comme source, la date à laquelle le document a été visionné sur notre site et de n'apporter aucune modification ni altération audit document, extrait ou illustration.

Vous trouverez à [cette page](#) récapitulées toutes les mentions légales qui encadrent la reproduction des documents parlementaires.

Cordialement,

La cellule internet du Sénat

Le 02/05/2021 à 10:31, Allouche Julie a écrit :

Monsieur Charles WALINE,

Je m'appelle Julie Allouche et je suis étudiante en master 2 d'éthique médicale et de bioéthique à l'Université de Paris (ex-Descartes).



Je rédige actuellement un mémoire de recherche sur le principe de garantie humaine, repris à l'article 11 du projet de loi de bioéthique, et votre schéma résumant les étapes de la discussion de la révision de cette loi est extrêmement clair !  
(<https://www.senat.fr/dossier-legislatif/pjl19-063.html>).

Serait-il possible, s'il vous plaît, de reproduire ce schéma dans mon mémoire pour expliquer le système de la navette parlementaire de cette loi ?

En vous remerciant par avance,  
Bien cordialement,  
Julie Allouche

| AVANT D'IMPRIMER, PENSEZ A L'ENVIRONNEMENT.  
|

## Autorisation 2 : Introduction du principe de Garantie Humaine à l'internationale, Health & Tech Intelligence – Care Insight, 01/06/2021.

**De:** Cécile DAVID [cecile.david@care-insight.fr](mailto:cecile.david@care-insight.fr)    
**Objet:** RE: Mémoire Garantie humaine - Webinar 25/05  
**Date:** 1 juin 2021 à 16:08  
**À:** Allouche Julie [julieall21@yahoo.fr](mailto:julieall21@yahoo.fr)  
**Cc:** Contact CARE INSIGHT [contact@care-insight.fr](mailto:contact@care-insight.fr), Sandrine Degos [sandrine.degos@care-insight.fr](mailto:sandrine.degos@care-insight.fr)

---

Bonjour Madame Allouche,

Tout d'abord merci à vous pour l'intérêt que vous portez à notre service Health & Tech Intelligence. Je confirme les propos de notre présidente, Sandrine Degos : aucun souci si vous nous citez (Health & Tech Intelligence - Care Insight).

Vous trouverez en PJ l'infographie en question.

Bon courage pour la finalisation de votre mémoire,

Bien cordialement,

**Cécile DAVID**  
Analyste-rédactrice, responsable éditoriale  
Health & Tech Intelligence

+33 6 33 38 36 59

8 rue Laugier - 75017 PARIS  
[www.care-insight.fr](http://www.care-insight.fr)  
[www.healthandtech.eu](http://www.healthandtech.eu)



---

**De :** Sandrine Degos  
**Envoyé :** mardi 1 juin 2021 15:37:50  
**À :** Allouche Julie; Cécile DAVID  
**Cc :** Contact CARE INSIGHT  
**Objet :** Re: Mémoire Garantie humaine - Webinar 25/05

bonjour Madame,  
pas de problème si vous citez la source. je vous mets en relation avec Cécile David qui en est l'auteur!

Envoyé de mon iPhone

> Le 1 juin 2021 à 14:29, Allouche Julie <[julieall21@yahoo.fr](mailto:julieall21@yahoo.fr)> a écrit :  
>  
> Madame Degos,  
>  
> Je m'appelle Julie Allouche et je suis étudiante en Master 2 d'éthique médicale et bioéthique à l'Université de Paris (ex-Descartes). Cette année, dans le cadre ce Master, je m'intéresse d'un point de vue éthique, au principe de garantie humaine du numérique en santé.  
>  
> Mes directeurs de mémoire, Madame le Professeur Mamzer et Monsieur Dupont, m'ont encouragé à assister à votre webinar du 25/05 à ce sujet. C'était très intéressant, merci !  
>  
> Votre schéma présenté au tout début de la conférence, récapitulant l'introduction du principe de garantie humaine à l'international (4 vignettes avec les drapeaux des pays) a éveillé mon intérêt : serait-il possible de l'intégrer à mon mémoire de recherche s'il vous plaît ? Si tel est le cas, pourriez