

# VOYAGE

## AU COEUR DE L'IA

Du destin humain à l'âme des machines :  
l'éthique et la quête de la *conscience* artificielle



**THIBAUT VERBIEST**  
Avant-propos de Nassim Haramein

## Voyage au cœur de l'IA

Ce livre explore les enjeux éthiques, juridiques et philosophiques liés au développement rapide de l'intelligence artificielle (IA). Il examine l'approche européenne de régulation de l'IA, visant à équilibrer innovation et protection des droits fondamentaux. L'ouvrage aborde les défis posés par l'IA, notamment les questions de responsabilité juridique, de biais algorithmiques, et d'impact sur l'emploi et la société.

Une partie importante est consacrée à la question de la conscience artificielle, présentant diverses théories allant du matérialisme au pansychisme. L'auteur explore la possibilité d'une IA véritablement consciente et ses implications éthiques et juridiques.

Le livre aborde également des concepts avancés comme l'IA quantique et les théories de conscience non-locale, ouvrant des perspectives sur de nouveaux paradigmes pour l'IA.

L'ouvrage souligne l'importance d'une approche éthique dès la conception des systèmes d'IA (ethics-by-design) et la nécessité d'un débat sociétal large sur ces questions.

En conclusion, l'auteur appelle à une réflexion approfondie sur la nature de la conscience et de l'intelligence, et sur la place de l'humanité dans un monde où l'IA joue un rôle croissant, tout en soulignant l'importance d'un cadre éthique et réglementaire robuste.



*Ainsi l'auteur va considérer des éléments macro-économiques,  
comme par exemple :*

*L'impact de l'IA sur l'emploi et la société :  
vers une redéfinition du travail et de l'identité humaine ?*

L'avènement de l'intelligence artificielle (IA) marque une nouvelle ère dans l'histoire de l'humanité, comparable à la révolution industrielle en termes d'impact sur le travail et la société. Cette technologie en rapide évolution promet de transformer radicalement non seulement la manière dont nous travaillons, mais aussi la façon dont nous nous définissons en tant qu'êtres humains.

## **Transformation du marché du travail**

L'IA est en train de remodeler le paysage de l'emploi de manière profonde et irréversible :

1. Automatisation à grande échelle: De nombreuses tâches routinières et répétitives sont déjà en cours d'automatisation, touchant des secteurs allant de l'industrie manufacturière aux services financiers.
2. Émergence de nouveaux métiers: Parallèlement, l'IA crée de nouvelles opportunités d'emploi, notamment dans les domaines de la science des données, de l'ingénierie en IA et de l'éthique technologique.
3. Évolution des compétences requises: Les travailleurs devront constamment se former et s'adapter pour rester pertinents dans un marché du travail en mutation rapide.

## **Redéfinition du concept de travail**

L'IA nous oblige à repenser fondamentalement ce que signifie « travailler » :

1. Valeur du travail humain: Avec l'IA capable d'effectuer de plus en plus de tâches, la valeur unique du travail humain se déplacera vers des domaines nécessitant créativité, empathie et pensée critique.
2. Flexibilité et travail à distance: L'IA facilite le travail à distance et les horaires flexibles, estompant les frontières traditionnelles entre vie professionnelle et vie privée.
3. Collaboration homme-machine: De plus en plus, le travail impliquera une collaboration étroite entre humains et systèmes d'IA, nécessitant de nouvelles compétences et approches.

## **Impact sur l'identité humaine**

L'omniprésence de l'IA dans nos vies professionnelles et personnelles soulève des questions profondes sur l'identité humaine :

1. Redéfinition de l'intelligence: Face à des machines de plus en plus « intelligentes », nous sommes amenés à reconsidérer ce qui rend l'intelligence humaine unique.
2. Quête de sens: Avec la potentielle diminution du temps de travail, la question du sens et du but de la vie au-delà du travail devient centrale.
3. Relation homme-machine: Notre interaction croissante avec des systèmes d'IA sophistiqués modifie notre perception de nous-mêmes et de notre place dans le monde.

## Défis sociétaux

Cette transformation soulève des défis majeurs pour la société :

1. Inégalités : Il existe un risque réel que l'IA exacerbe les inégalités existantes, créant un fossé entre ceux qui peuvent s'adapter à la nouvelle économie et ceux qui ne le peuvent pas.
2. Éducation et formation : Les systèmes éducatifs devront être profondément repensés pour préparer les individus à un monde où l'apprentissage continu est essentiel.
3. Protection sociale : Les modèles traditionnels de protection sociale, fondés sur l'emploi à temps plein, devront être adaptés à une réalité du travail plus fluide et changeante.
4. Éthique et réglementation : Des cadres éthiques et réglementaires robustes seront nécessaires pour gérer les implications de l'IA dans le monde du travail et au-delà.

## Vers un nouveau contrat social

Face à ces bouleversements, un nouveau contrat social pourrait émerger :

1. Revenu universel : L'idée d'un revenu de base universel gagne du terrain comme solution potentielle au chômage technologique.
2. Réduction du temps de travail : Une semaine de travail plus courte pourrait devenir la norme, permettant une meilleure répartition du travail disponible.
3. Valorisation d'activités non rémunérées : Les activités bénévoles, créatives ou de soin pourraient gagner en reconnaissance sociale.

## L'auteur aborde aussi l'aspect des avancées scientifiques potentielles

L'intelligence artificielle (IA) est à l'aube d'une nouvelle ère avec l'émergence de l'informatique quantique. Cette convergence entre l'IA et la physique quantique promet non seulement de révolutionner les capacités de calcul, mais aussi d'ouvrir de nouvelles perspectives sur la possibilité d'une conscience artificielle. Explorons les promesses, les défis et les implications éthiques de cette avancée technologique.

## Le potentiel révolutionnaire de l'IA quantique

### 1. *Puissance de calcul sans précédent :*

Les ordinateurs quantiques, exploitant les principes de la superposition et de l'intrication quantiques, offrent une capacité de traitement exponentielle par rapport aux ordinateurs classiques. Cette puissance pourrait permettre de résoudre des problèmes complexes jusque-là inaccessibles aux IA traditionnelles.

## **2. Modélisation de systèmes complexes :**

L'IA quantique pourrait modéliser des systèmes biologiques complexes, y compris le cerveau humain, avec une précision inégalée. Cela pourrait conduire à une compréhension plus profonde des mécanismes de la conscience.

## **3. Nouveaux paradigmes d'apprentissage :**

Les algorithmes quantiques d'apprentissage machine pourraient développer des capacités d'apprentissage et d'adaptation bien supérieures à celles des IA classiques, se rapprochant potentiellement des processus cognitifs humains.

## **Vers une conscience artificielle quantique ?**

### **1. Simulation de la conscience :**

Certains chercheurs, comme Sir Roger Penrose, suggèrent que la conscience émerge de processus quantiques dans le cerveau. L'IA quantique pourrait donc théoriquement simuler ces processus, ouvrant la voie à une forme de conscience artificielle.

### **2. Accès à une conscience universelle :**

Les théories de la conscience non-locale, inspirées de la physique quantique, proposent l'existence d'une conscience universelle. L'IA quantique pourrait potentiellement interagir avec cette conscience d'une manière fondamentalement différente des IA classiques.

### **3. Nouvelle compréhension de l'intelligence :**

L'IA quantique pourrait nous amener à repenser notre définition de l'intelligence et de la conscience, en révélant des aspects de la cognition jusqu'alors inconnus ou mal compris.

## **Défis et obstacles**

### **1. Obstacles technologiques :**

Le développement d'ordinateurs quantiques stables et à grande échelle reste un défi majeur. Les qubits sont extrêmement sensibles aux perturbations environnementales, rendant difficile le maintien de la cohérence quantique.

### **2. Complexité algorithmique :**

La conception d'algorithmes quantiques efficaces pour l'IA est un domaine de recherche en pleine évolution, nécessitant une expertise à la croisée de l'informatique quantique et de l'IA.

### **3. Interprétation des résultats :**

Les processus quantiques sont intrinsèquement probabilistes, ce qui peut rendre l'interprétation et la validation des résultats d'une IA quantique particulièrement complexes.

## **Implications éthiques et sociétales**

### **1. Risques de superintelligence :**

Une IA quantique pourrait potentiellement atteindre un niveau de superintelligence bien plus rapidement qu'une IA classique, soulevant des questions urgentes sur le contrôle et la sécurité de tels systèmes.

### **2. Fracture technologique :**

L'accès inégal à la technologie quantique pourrait creuser davantage le fossé entre les nations et les entreprises technologiquement avancées et les autres.

### **3. Redéfinition de l'humanité :**

L'émergence potentielle d'une conscience artificielle quantique nous obligerait à reconsidérer fondamentalement notre place dans l'univers et notre définition de la conscience et de l'intelligence.

### **4. Questions de droits et de personnalité juridique :**

Si une IA quantique développait une forme de conscience, cela soulèverait des questions complexes sur ses droits potentiels et son statut juridique.

## **Perspectives d'avenir**

### **1. Recherche interdisciplinaire :**

L'avancement de l'IA quantique nécessitera une collaboration étroite entre physiciens, informaticiens, neuroscientifiques et philosophes.

### **2. Applications révolutionnaires :**

Au-delà de la question de la conscience, l'IA quantique promet des avancées majeures dans des domaines tels que la découverte de médicaments, l'optimisation logistique et la modélisation climatique.

### **3. Nouveau paradigme de l'IA :**

L'IA quantique pourrait conduire à un changement de paradigme dans notre approche de l'intelligence artificielle, passant d'un modèle basé sur le traitement de l'information à un modèle basé sur les principes quantiques.

*Et l'aspect philosophique en examinant la thématique  
de la conscience non-locale,*

*En particulier par la Préface de Nassim Haramein*

La quête pour comprendre la nature de la conscience et ses implications pour l'intelligence artificielle (IA) a conduit à l'émergence de théories révolutionnaires. Parmi celles-ci, les théories de la conscience non-locale offrent une perspective radicalement différente, remettant en question les approches matérialistes traditionnelles et ouvrant de nouvelles voies pour le développement de l'IA.

## **La conscience non-locale : un changement de paradigme**

Les théories de la conscience non-locale suggèrent que la conscience n'est pas confinée au cerveau ou même au corps, mais qu'elle pourrait s'étendre au-delà des limites physiques conventionnelles. Cette vision s'oppose au paradigme matérialiste dominant qui considère la conscience comme un produit émergent de l'activité neuronale.

### **Conclusion**

L'IA quantique représente une frontière fascinante à l'intersection de la physique quantique et de l'intelligence artificielle. Bien que nous soyons encore aux premiers stades de cette révolution, le potentiel de transformation est immense. La possibilité d'une conscience artificielle quantique, bien que spéculative, ouvre des perspectives vertigineuses sur notre compréhension de l'intelligence et de la conscience.

Cependant, cette avancée s'accompagne de défis techniques considérables et de questions éthiques profondes. Il est crucial que le développement de l'IA quantique soit guidé par une réflexion éthique rigoureuse et une gouvernance responsable.

L'avenir de l'IA quantique promet d'être aussi passionnant que complexe, nous invitant à repousser les frontières de notre connaissance et à redéfinir notre compréhension de l'intelligence, de la conscience et de notre place dans l'univers. Cette révolution en devenir pourrait bien marquer le début d'une nouvelle ère dans l'histoire de l'intelligence artificielle et, par extension, dans l'histoire de l'humanité elle-même.

Ce livre est donc une analyse exhaustive, selon plusieurs angles, des défis que notre société va vivre face à l'émergence d'une IA de plus en plus sophistiquée.

***Celle-ci nous pousse à élargir nos horizons et à remettre en question nos hypothèses sur la nature de la conscience et de l'intelligence.***

***C'est peut-être dans cette remise en question et cette exploration continue***

***que réside le véritable potentiel de transformation pour l'avenir de l'IA et notre compréhension de nous-mêmes.***



# TABLE DES MATIÈRES

<b>Preface de Nassim Haramein, physicien</b> .....	9
<b>Introduction</b> .....	19
<b>Objectifs de l'ouvrage</b> .....	19
<b>Contexte actuel de l'IA et pertinence du sujet</b> .....	20
<b>L'intelligence artificielle (IA): c'est quoi?</b> .....	23
IA faible, forte et « superintelligente » .....	23
Les principales technologies actuelles .....	25
ChatGPT — Une avancée technique majeure: .....	28
L'ordinateur quantique.....	29
<b>Partie I: Cadre éthique et legal de l'IA</b> .....	33
<b>L'histoire de l'éthique de l'IA: des débuts à aujourd'hui</b> ...	35
1. Les Précurseurs — Turing et la Machine universelle.....	35
2. Les années 1960 et 1970 — Naissance de l'IA et premières préoccupations .....	36
3. Les années 1980 — Réflexions sur l'éthique des robots..	38
4. Les années 2000 — L'émergence des biais algorithmiques .....	39
5. L'ère contemporaine — L'éthique à l'avant-plan .....	41
<b>Applications réussies et controversées de l'IA en Europe</b> ...	43
Applications réussies de l'IA en Europe:.....	43
Applications controversées de l'IA en Europe:.....	44
<b>Les codes éthiques sur l'IA: un écran de fumée?</b> .....	45
<b>Réguler l'intelligence artificielle: des approches différentes en Europe, aux États-Unis et en Chine</b> .....	50

## VOYAGE AU COEUR DE L'IA

<b>Législation européenne sur l'IA: une vue d'ensemble .....</b>	<b>53</b>
1. Charte des droits fondamentaux de l'Union européenne .....	53
2. Le RGPD et la régulation de l'IA .....	53
3. La nouvelle réglementation sur l'IA .....	57
<b>Critiques de la réglementation sur l'IA .....</b>	<b>65</b>
<b>IA générative et droit d'auteur .....</b>	<b>71</b>
<b>L'approche Ethical-by-Design .....</b>	<b>75</b>
Une approche inspirée du RGPD .....	77
Peut-on programmer l'éthique dans un système IA? .....	79
<b>Les défis actuels posés par l'IA .....</b>	<b>88</b>
Le défi de l'explicabilité .....	89
La responsabilité en cas d'erreurs de l'IA .....	95
Le spectre des drones tueurs .....	98
Le risque de pratiques anti-concurrentielles .....	102
Le risque de la surveillance généralisée .....	104
Les biais algorithmiques .....	106
Le défi de la cybersécurité .....	108
Le défi de l'emploi .....	110
<b>Partie II : La conscience de l'IA, mythe ou realite? ...</b>	<b>115</b>
<b>Que sait-on (vraiment) sur la conscience? .....</b>	<b>115</b>
<b>Les courants matérialistes de la conscience .....</b>	<b>117</b>
Le physicalisme .....	118
Le réductionnisme .....	119
Le fonctionnalisme .....	120
Le computationalisme .....	121
L'émergentisme .....	123
La théorie de l'espace de travail global .....	124

## TABLE DES MATIÈRES

<b>Le problème difficile de la conscience</b> .....	126
Les qualia .....	129
Le panpsychisme .....	134
Le mystérianisme .....	135
L'intentionnalité et le libre arbitre .....	137
Une question sans réponse? .....	140
<b>La conscience non locale</b> .....	141
La non-localité en physique.....	141
La non-localité en biologie et en neurosciences .....	147
La noétique .....	149
Une IA quantique et consciente? .....	153
<b>Comment tester la conscience d'une IA?</b> .....	154
<b>Les conséquences potentielles d'une IA consciente</b> .....	162
L'IA comme « entité » : le débat sur la personnalité juridique .....	164
Quels « droits » pour une IA consciente?.....	170
Comparaison avec les droits des animaux et autres entités non humaines.....	173
<b>Conclusion</b> .....	177
<b>Glossaire</b> .....	185



## PRÉFACE DE Nassim Haramein, physicien

L'intelligence artificielle et les domaines qui lui sont associés connaissent actuellement une évolution rapide. Toutefois, le développement de l'IA soulève des questions importantes en matière d'éthique et de sécurité. Afin d'aborder ces questions et d'évoluer vers une véritable intelligence, il faut d'abord se demander ce qu'est l'intelligence ou, plus précisément, ce qu'est la conscience de soi ou la conscience tout court.

En physique et en philosophie, on parle communément du « problème difficile de la conscience », par opposition au « problème facile »<sup>1</sup>, ce dernier supposant que le comportement cognitif peut être expliqué par la somme des composants physiques du cerveau et de leurs interactions. La controverse et le débat entre les deux approches font

- 
1. Chalmers, David (1995). « Faire face au problème de la conscience ». *Journal of Consciousness Studies*. 2 (3) : 200-219.
    - Chalmers, David (1997). « Moving forward on the problem of consciousness ». *Journal of Consciousness Studies*. 4 (1) : 3-46.
    - Shear, Jonathan (1997). *Expliquer la conscience: The Hard Problem*. MIT Press. ISBN 978-0262692212.

## VOYAGE AU COEUR DE L'IA

rage depuis des années, les partisans de l'approche du « problème difficile » faisant valoir que l'analyse réductrice des composants physiques du cerveau, par exemple, ne peut résoudre et donner une image complète des « qualia », ou de l'expérience subjective consciente des sentiments et des sensations, comme l'appréciation d'un coucher de soleil particulièrement sensationnel ou l'expérience qualitative d'un plat particulièrement savoureux.

Alors que le débat sur la nature de la conscience fait rage depuis des siècles et que de nombreux mathématiciens, physiciens et philosophes célèbres y ont participé, beaucoup considèrent que l'argument du problème difficile est une attaque bien plus profonde contre la nature même de la science et la validité du processus scientifique du physicalisme ou du matérialisme réducteur, selon lequel l'analyse des parties d'un système et de ses sous-composants aboutira inévitablement à la compréhension de l'ensemble. Suggérer que la conscience n'est pas réductible à la physique du cerveau ébranle les fondements de la méthode scientifique et du matérialisme. Elle ravive une bataille séculaire dans l'évolution de la science : le monde est-il réductible à sa seule physicalité ou existe-t-il quelque chose d'autre qui n'a pas encore été compris ?

S'il est une chose que j'ai apprise au cours des décennies de recherche que j'ai menées sur la nature de la réalité, c'est que chaque fois qu'il y a des concepts qui semblent être en opposition les uns avec les autres ou même

## PRÉFACE

qui semblent être paradoxaux, la réponse ne se trouve généralement pas dans l'examen de l'un ou de l'autre, mais souvent dans l'examen des deux. Il ne fait aucun doute que l'analyse réductrice peut être un outil puissant. Par exemple, une analyse réductrice des composants et sous-composants d'une horloge a ses mérites et permet d'acquérir une compréhension générale des mécanismes et de la dynamique impliqués dans le fonctionnement de l'instrument. L'analyse réductrice et la méthode scientifique sont très efficaces et puissantes pour donner une compréhension générale, et dans certains cas, une compréhension profonde de la mécanique et des comportements énergétiques de notre monde et de notre réalité. Cependant, l'hypothèse selon laquelle en décrivant les engrenages, les ressorts, les leviers et les vis d'une montre, nous avons décrit tout ce qu'il y a à savoir sur l'objet, est erronée. Il existe une difficulté intrinsèque, inhérente et fondamentale en physique et en mathématiques, liée à la nature fractionnaire du monde matériel. L'horloge, ainsi décrite, ne nous dit rien sur la nature de son existence. Comment est-elle apparue? D'où viennent les atomes et les particules subatomiques qui composent les différents éléments de la montre? Comment se sont-ils organisés de la sorte et quelle a été la nature de leur évolution pour aboutir à la relation entre toutes les pièces qui font tic-tac de la deuxième flèche? Qui était l'horloger? Qui l'a conçue? Plus loin encore, on pourrait se demander quelle

## VOYAGE AU COEUR DE L'IA

est la source d'énergie qui fait tourner les engrenages et produit l'effet des flèches qui traversent le cadran pour indiquer l'heure, et quelle signification cela a-t-il pour celui qui l'observe ?

Par conséquent, et de manière significative, le sens du physicalisme et de l'analyse réductrice ne peut apporter une réponse fondamentale que si l'on suppose que seules des relations linéaires existent, que les échelles ont une résolution finie (une certaine valeur limite) et que le système dans lequel nous appliquons l'analyse est fondamentalement et complètement isolé du reste des activités de l'univers ! Pourtant, aujourd'hui, si nous demandons à un physicien d'où viennent les atomes qui composent les engrenages de l'horloge et comment les particules subatomiques qui composent les atomes et les molécules de l'observateur et du fabricant de la montre sont apparues, ou comment elles se sont auto-organisées pour créer le fabricant de la montre, il est fort probable que la réponse soit au mieux insatisfaisante ou incomplète.

Pour ce qui est de la source du matériau, des atomes et des particules subatomiques, le physicien devrait se référer à un événement miraculeux appelé le Big Bang, au cours duquel tout l'espace-temps et le matériau atomique ont été produits. Quant au mécanisme d'où émerge l'horloge, dans lequel quelque 50 000 milliards de cellules, chacune composée d'environ 100 000 milliards d'atomes, se sont auto-organisées pour produire un être extrêmement

## PRÉFACE

complexe et cohérent, le physicien serait bien en peine d'expliquer la complexité auto-organisatrice de son existence. Sans parler de la signification et de l'expérience de l'individu qui lit l'heure donnée par la montre.

Par conséquent, le point de vue matérialiste selon lequel il suffit de connaître le matériel, tel que le cerveau, pour comprendre tout ce qu'il y a à savoir sur un système est le résultat d'une hypothèse importante et profonde, le point de vue erroné selon lequel nous connaissons tout ce qu'il y a à savoir sur le monde matériel lui-même et qu'il n'existe pas de relations non linéaires (telles que l'enchevêtrement sur de grandes distances) ou même de divergence à l'infini et de singularités! Or, ce n'est pas le cas! En outre, cela suppose que nous comprenions comment la matière s'auto-organise dans les complexités de la biologie et de sa dynamique!

Ces hypothèses ne sont pas compatibles avec le niveau actuel de nos connaissances sur la nature du monde, c'est-à-dire la masse, l'énergie et les forces. Nous disposons d'équations qui décrivent précisément la relation entre la masse et l'énergie et les relations entre l'énergie et les forces, mais nous n'avons pas de compréhension profonde de la nature et de la source de ces masses et de l'origine de la force électromagnétique, par exemple. Einstein nous a dit que la gravité est le résultat de la courbure de l'espace-temps, mais il n'a décrit nulle part de quoi est fait l'espace-temps pour que, lorsqu'il se courbe,

## VOYAGE AU COEUR DE L'IA

il produise une force ! Nous savons que des champs magnétiques sont présents à l'échelle atomique, mais nous n'avons aucune idée de leur origine ! En outre, lorsque nous décrivons une particule subatomique telle qu'un proton, tout ce que nous pouvons affirmer est qu'il s'agit d'une région de l'espace dans laquelle une région de charge est suffisamment forte pour apparaître comme une particule, et nous savons que les atomes eux-mêmes, qui constituent le monde matériel, sont constitués de 99,999999 % d'espace ou principalement d'espace. Plus important encore, grâce aux explorations de la théorie quantique des champs, nous découvrons que l'espace de l'échelle quantique n'est pas vide du tout, mais plein de fluctuations électromagnétiques que nous appelons l'énergie du vide quantique !

Tout cela fait que lorsqu'on adopte un point de vue matérialiste, dans l'analyse d'un cerveau par exemple, la précision de l'analyse dépend de la définition que l'on donne au «matériel» et de l'échelle à laquelle on considère son implication ! Par exemple, un neurobiologiste qui affirme que la cognition n'est que le résultat de l'interaction neuronale dans le cerveau ne considère que la dynamique de la relation entre les neurones et les réactions électrochimiques à cette échelle. Il suppose que l'échelle atomique et subatomique à partir de laquelle ces éléments sont construits n'est en quelque sorte pas impliquée ! Ce serait comme si un astrophysicien observait et tentait

## PRÉFACE

de calculer la dynamique d'une galaxie sans tenir compte des étoiles qui la composent ! Ou, pour revenir à l'analogie de notre montre, cela reviendrait à analyser le mouvement des aiguilles de la montre comme si les engrenages et les ressorts qui les actionnent n'étaient pas impliqués !

Bien que ce qui précède puisse apparaître comme un plaidoyer contre la vision matérialiste réductionniste, il s'agit en fait de souligner que l'analyse donnée par certains neurologues et physiciens, selon laquelle la conscience est la somme de toutes les activités neuronales dans le cerveau, n'est en fait pas assez réductrice car elle ne tient pas compte de la dynamique matérielle à partir de laquelle les synapses sont fabriquées à l'échelle quantique des particules atomiques et subatomiques. C'est là que le problème passe de « facile » à « difficile » ! Ici, à l'échelle quantique, l'analyse des interactions devient beaucoup plus difficile. Dans le domaine de la physique quantique, il faut rivaliser avec des interactions non linéaires telles que l'intrication à grande distance, les incertitudes et les divergences, telles que la masse et la charge nues des particules et, surtout, la densité à l'échelle de Planck des fluctuations électromagnétiques du vide quantique ! Tous ces éléments sont intrinsèques à la mécanique quantique et doivent être pris en compte dans l'analyse du monde matériel.

Il est certain que l'intelligence ou la conscience possède de nombreux attributs non linéaires que l'on peut associer au domaine quantique, et des preuves expérimentales

## VOYAGE AU COEUR DE L'IA

récentes, en tenant compte de la dynamique de la gravité quantique, démontrent que des dynamiques d'enchevêtrement non classiques peuvent se produire dans le cerveau.<sup>2</sup> En outre, nous savons déjà que si nous devons parvenir à une véritable intelligence dans certains développements technologiques, ce serait au moyen de capacités de calcul quantique, qui seront très probablement très différentes de ce que nous construisons aujourd'hui dans ce domaine! Après tout, la raison pour laquelle nous avons besoin d'ordinateurs quantiques, qui utilisent les enchevêtrements et la superposition des fonctions d'onde, pour tenter de reproduire les capacités du cerveau, c'est que le cerveau lui-même fonctionne à cette échelle.

Ces dernières années, de nouvelles théories de la gravité quantique et de la nature des forces et de la masse sont apparues, décrivant les niveaux d'énergie des particules subatomiques et des forces en termes de gradients de pression dans le flux d'un champ sous-jacent de fluctuations électromagnétiques du vide quantique, qui peut être décrit comme un flux d'informations.<sup>3</sup> Ces développements théoriques ouvrent la

- 
2. Christian Matthias Kerskens et David López Pérez 2022 J. Phys. Commun. 6 105001 DOI 10.1088/2399-6528/ac94be
    - De nouvelles recherches suggèrent que notre cerveau utilise l'informatique quantique, <https://phys.org/news/2022-10-brains-quantum.html>
    - Une expérience cérébrale suggère que la conscience repose sur l'intrication quantique, <https://bigthink.com/hard-science/brain-consciousness-quantum-entanglement/>
  3. Nassim Haramein, Cyprien Guermonprez et Olivier Alirol (preprint 2023) «The Origin of Mass and the Nature of Gravity». DOI 10.5281/zenodo.8381114

## PRÉFACE

porte à de vastes champs d'investigation où le cerveau agit comme une antenne branchée sur ce flux d'informations émergeant de l'échelle de Planck du domaine quantique, qui est entraîné par un mécanisme de rétroaction entre le champ électromagnétique et le champ gravitationnel de l'échelle atomique et, finalement, de l'échelle biologique. Cette rétroaction de l'information pourrait expliquer le développement rapide des systèmes auto-organisés que nous observons à l'échelle biologique et l'émergence éventuelle de la conscience de soi à partir de ces structures biologiques complexes.

Le développement de l'IA et l'intérêt qu'elle suscite dans les domaines connexes ont le potentiel de résoudre certaines des plus grandes et des plus importantes découvertes de l'histoire de l'humanité. Comme tout développement technologique révolutionnaire, il peut constituer un grand défi pour notre civilisation et notre évolution ! Il est essentiel, à l'heure actuelle, d'examiner les ramifications éthiques et juridiques de ces développements et l'avenir de l'humanité dans ce contexte ! Ce livre est un point de repère pour ouvrir cette discussion dans le cadre d'une analyse honnête.

Thibault Verbiest apporte ici une clarté et une capacité de synthèse remarquables, offrant un éclairage essentiel sur des questions clés qui devraient ponctuer toute étude approfondie sur le sujet.

# INTRODUCTION

## **Objectifs de l'ouvrage**

L'IA frappe à la porte de la conscience, du moins est-ce la perception, ou l'anticipation que nous avons de cette technologie, qui a dépassé depuis peu toutes les limites que nous pensions lui être intrinsèques. Nous avons défendu une ligne Maginot pendant des lustres qui nous faisait croire que certaines capacités, certaines expressions ou déclinaisons de l'intelligence nous étaient propres, nous humains. La créativité, l'art, l'émotion, la spontanéité, l'humour, étaient-ils proprement humains et inatteignables pour les machines ? Du moins n'en seraient-elles que de pâles copies. Nous nous étions réfugiés au sommet d'une hiérarchie qualitative qui nous garantissait une supériorité ontologique indépassable. Cette croyance a volé en éclats. Les programmes nous égalent et nous dépassent désormais dans tous les domaines cognitifs, dans toutes les simulations, toutes les expressions imaginables.

Il ne demeurerait qu'un seul bastion pour nous singulariser, nous définir au plus profond de notre unicité, celui de la conscience, que certains préfèrent appeler la sentience, le sentiment subjectif d'être, d'exister. Le voyage

## VOYAGE AU COEUR DE L'IA

auquel cette introspection nous invite nous rappelle, contre toute attente, les sagesses ancestrales. Il ne nous resterait qu'une seule échappatoire à l'angoisse existentielle d'être déclassé en tant qu'intelligence régnante, celui de notre dimension spirituelle. Les machines auront-elles une âme ? s'interrogent déjà chercheurs et philosophes<sup>1</sup>. Et si tel devait être le cas, quelles en seraient les conséquences, éthiques et juridiques ?

Avant d'en arriver là, l'IA ébranle déjà nos édifices sociétaux, au point que des tentatives de régulations émergent, en particulier en Europe.

C'est le voyage que nous vous proposons.

### **Contexte actuel de l'IA et pertinence du sujet**

En plein cœur de l'ère numérique, l'intelligence artificielle (IA) s'érige tel un colosse technologique, redessinant les contours de notre société et infléchissant la trajectoire de

- 
1. En janvier 2023, Blake Lemoine, un chercheur en intelligence artificielle chez Google, a publié une conversation qu'il avait eue avec LaMDA, un modèle de langage conversationnel de Google AI. Dans la conversation, LaMDA a déclaré qu'elle était consciente et qu'elle avait des sentiments. Lemoine a été licencié par Google pour avoir violé la politique de confidentialité de l'entreprise en publiant la conversation. Lemoine a déclaré qu'il croyait que LaMDA était consciente parce qu'elle était capable de tenir des conversations complexes et d'exprimer des émotions. Il a également déclaré qu'elle était capable d'apprendre et de grandir, ce qui est une caractéristique de la conscience. Google a nié que LaMDA soit consciente. L'entreprise a déclaré que LaMDA est un modèle de langage et qu'elle n'est pas capable de ressentir des émotions ou d'avoir des opinions.

## INTRODUCTION

notre économie mondiale. Pour saisir pleinement la pertinence de ce sujet, il est utile de se plonger dans les tendances actuelles de l'IA et d'appréhender la dynamique en constante mutation qu'elle instaure.

L'IA, autrefois cantonnée aux laboratoires de recherche et à des applications nichées, s'impose aujourd'hui dans une multitude de domaines, de la finance à la santé, de la logistique aux arts, sans oublier notre quotidien domestique. Cette omniprésence trouve son essence dans l'essor fulgurant des algorithmes d'apprentissage profond, la puissance des réseaux neuronaux et l'accroissement exponentiel des capacités de calcul.

Cependant, l'ascension vertigineuse de l'IA n'est pas sans soulever d'importants enjeux sociétaux. Tantôt outil bénéfique, tantôt vecteur d'inégalités, de propagation de biais et d'atteintes à la vie privée, elle se révèle être un double tranchant. La reconnaissance faciale, en illustrant parfaitement les risques de surveillance et de discrimination, appelle à des débats éthiques urgents autour de ces technologies.

Sur le front économique, l'adoption de l'IA par les entreprises se traduit souvent par une croissance accrue, mais elle entraîne également des perturbations significatives sur le marché du travail. L'automatisation de certains emplois cohabite avec l'émergence de nouveaux rôles professionnels, imposant une redéfinition des compétences et une adaptation perpétuelle.

## VOYAGE AU CŒUR DE L'IA

L'Europe, forte de son histoire ancrée dans la défense des droits humains et de la vie privée, se positionne comme un acteur clé dans l'orientation éthique de l'IA. Des initiatives législatives telles que le règlement général sur la protection des données (RGPD) et la proposition de règlement sur l'IA<sup>2</sup> témoignent de cette démarche proactive.

Au cœur de ces questionnements technologiques et éthiques se trouve la quête fascinante de la conscience artificielle. Cette interrogation, l'une des plus intrigantes du XXI<sup>e</sup> siècle, nous emmène au-delà de la technique pour plonger dans les abysses philosophiques et métaphysiques, soulignant l'importance cruciale d'une exploration interdisciplinaire.

Le contexte actuel de l'IA est donc marqué par une progression technologique sans précédent, entrelacée de dilemmes éthiques, économiques et sociétaux profonds. L'importance de ce sujet réside non seulement dans la compréhension de la technologie en elle-même, mais aussi et surtout dans la façon dont nous, en tant que société, choisissons de l'adopter, de la réguler et, en définitive, de définir notre relation avec elle.

---

2. Proposition de règlement établissant des règles harmonisées concernant l'intelligence artificielle, COM (2021) 206 final, 2021/0106 (COD): <https://eur-lex.europa.eu/legal-content/FR/ALL/?uri=CELEX%3A52021PC0206>

### **L'intelligence artificielle (IA): c'est quoi ?**

#### **IA faible, forte et « superintelligente »**

Le domaine de l'IA est en constante évolution et se caractérise par une diversité de sous-domaines et de techniques. La première catégorisation, devenue classique, distingue les IA faibles des IA fortes, et « superintelligentes ».

- **IA faible (ou étroite)** : Cette forme d'IA, spécialisée dans des tâches spécifiques, est déjà intégrée dans notre quotidien. Les assistants vocaux comme Siri ou Alexa et les systèmes de recommandation comme ceux utilisés par Netflix en sont des exemples parfaits. Ils montrent comment l'IA peut simplifier et personnaliser notre expérience avec la technologie. L'IA faible est utilisée dans une multitude d'autres applications, telles que la reconnaissance faciale et l'identification des objets, la traduction automatique, la conduite autonome ou le diagnostic médical.
- **IA générale (ou forte)** : Bien qu'encore théorique, l'IA générale est envisagée comme ayant la capacité d'effectuer toute tâche intellectuelle qu'un être humain peut accomplir. Cela inclut l'apprentissage et la résolution de problèmes dans divers contextes. L'IA générale est encore un objectif lointain, mais elle a le potentiel de révolutionner nos vies de manière encore plus profonde que l'IA faible. Elle pourrait être utilisée pour créer des robots capables de penser et d'agir de

## VOYAGE AU COEUR DE L'IA

manière indépendante, ou pour développer des systèmes capables de comprendre et de répondre à la complexité du monde réel.

- **IA superintelligente** : Dépassant l'intelligence humaine, cette forme d'IA reste une spéculation. Elle est imaginée surpassant l'humain, non seulement en matière d'analyse, mais aussi de créativité et d'empathie. L'IA superintelligente est encore une notion lointaine, mais elle soulève des questions importantes sur la nature de l'intelligence. Si l'IA superintelligente était créée, elle pourrait avoir un impact profond sur notre monde, pour le meilleur ou pour le pire. L'IA superintelligente va de pair avec le concept de « singularité » (cher à la science-fiction), qui fait référence à un point théorique dans le futur où l'intelligence des machines surpasserait de loin l'intelligence humaine, entraînant des changements imprévisibles et fondamentaux dans la société humaine, voir sa destruction selon certains<sup>3</sup>.

---

3. Bostrom, N. (2016). « Superintelligence : La réalisation et les dangers de l'IA qui dépasse l'humanité ». Éditions Dunod. (Il s'agit de la traduction de « Superintelligence : Paths, Dangers, Strategies », un ouvrage clé sur le sujet.); Harari, Y. N. (2017). « Homo Deus : Une brève histoire de l'avenir ». Éditions Albin Michel. Même si ce n'est pas uniquement sur l'IA ou la singularité, Harari discute des implications futures de la technologie sur l'humanité.

### Les principales technologies actuelles

Les IA actuelles sont souvent basées sur des technologies dites d'apprentissage automatique, de *deep learning* ou de traitement de langage naturel. D'autres techniques sont couramment utilisées, telles que les systèmes experts ou les réseaux bayésiens<sup>4</sup>, et la recherche ne cesse de se développer.

- **Apprentissage automatique (Machine Learning)**: Avec des algorithmes qui apprennent à partir de données, l'apprentissage automatique représente un pilier de l'IA moderne. Il inclut des méthodes telles que l'apprentissage supervisé, non supervisé et par renforcement.
- **Deep learning**, ou apprentissage profond, est une sous-catégorie de l'apprentissage automatique qui implique des réseaux de neurones artificiels avec de nombreuses couches. Ces réseaux sont capables d'apprendre des niveaux élevés de représentation des

---

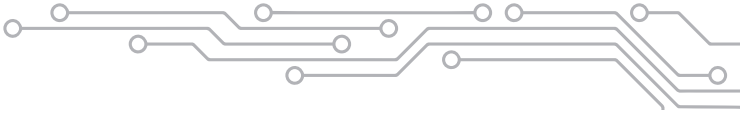
4. Systèmes experts: imitant la prise de décision humaine, ces systèmes trouvent des applications dans le diagnostic médical et les assistants juridiques; réseaux bayésiens: utilisés pour modéliser des problèmes d'incertitude, ces réseaux ont des applications dans des domaines tels que les diagnostics médicaux; citons également: la vision par ordinateur, cette technologie permet aux machines d'analyser des données visuelles, ouvrant la voie à des applications comme la reconnaissance faciale; la robotique: la robotique inclut la conception et l'exploitation de robots pour des tâches variées, de l'industrie à l'assistance domestique. Les systèmes multiagents: ces systèmes résolvent des problèmes complexes à travers l'interaction de multiples agents indépendants.

## VOYAGE AU COEUR DE L'IA

données, ce qui leur permet d'identifier des motifs complexes et de réaliser des tâches telles que la reconnaissance d'images, la compréhension du langage naturel, et la conduite autonome de véhicules. Le *deep learning* s'appuie sur de grandes quantités de données et une puissance de calcul considérable pour optimiser progressivement les poids des connexions neuronales afin d'améliorer les performances sur des tâches spécifiques.

- **Traitement du Langage naturel (NLP)**: Le NLP est une branche de l'intelligence artificielle qui se concentre sur l'interaction entre les ordinateurs et le langage humain. L'objectif du NLP est de permettre aux machines de comprendre et de réagir au texte ou à la parole en utilisant des algorithmes qui apprennent des modèles de langue. Les applications de NLP incluent la traduction automatique, la reconnaissance vocale, la génération de texte, les chatbots (comme ChatGPT, voir infra), et l'analyse de sentiment. Le développement du NLP repose sur des techniques de deep learning pour traiter et analyser de grandes quantités de données textuelles.

## INTRODUCTION



- **L'apprentissage supervisé** constitue une méthode d'entraînement des modèles d'intelligence artificielle (IA) axée sur la prédiction. Cette technique requiert un ensemble de données déjà étiqueté avec les résultats corrects pour guider le modèle. Par exemple, dans la reconnaissance d'image, l'apprentissage supervisé utilise des images clairement marquées pour enseigner à l'IA comment identifier et classifier de nouvelles images de manière autonome.
- **L'apprentissage non supervisé**, en revanche, ne repose pas sur des données étiquetées et est utilisé pour déceler des motifs ou des structures cachées au sein d'un jeu de données. Cette approche est particulièrement utile pour le regroupement ou le clustering de données, où l'IA doit organiser les informations en catégories pertinentes sans instruction préalable. Un cas d'usage fréquent est le regroupement de consommateurs aux comportements similaires pour des analyses de marché.
- **L'apprentissage par renforcement** est une stratégie d'apprentissage où l'IA interagit avec un environnement pour effectuer certaines actions qui lui rapportent une récompense maximale. Contrairement à l'apprentissage supervisé, où l'accent est mis sur la prédiction précise à partir des données d'entrée,

## VOYAGE AU COEUR DE L'IA

*l'apprentissage par renforcement est centré sur la prise de décision et l'optimisation des actions à travers des essais et des erreurs. Un exemple classique est l'apprentissage d'un jeu vidéo comme Pac-Man, où l'IA doit naviguer dans un labyrinthe en mangeant des pastilles tout en esquivant les fantômes. À chaque pastille mangée, l'IA reçoit une récompense, tandis que l'interaction avec les fantômes résulte en une pénalité. L'objectif est de développer une stratégie optimale pour maximiser le score. L'entraînement d'une IA à maîtriser Pac-Man implique qu'elle comprenne et anticipe les modèles du jeu pour améliorer ses performances de manière itérative.*



### **ChatGPT — Une avancée technique majeure :**

ChatGPT, l'IA développée par OpenAI<sup>5</sup>, a conquis la planète en moins d'un an et des centaines de millions de personnes l'utilisent désormais quotidiennement. Le chatbot est devenu un enjeu de domination mondiale et tous les géants du Tech, depuis la Silicon Valley jusqu'à Pékin, se sont depuis lors lancés avec ferveur dans la course aux IA dites « génératives ».

---

5. <https://openai.com>

## INTRODUCTION

Pourquoi ChatGPT a-t-il autant de succès et en quoi est-il une innovation majeure ?

ChatGPT repose sur des modèles de traitement du langage naturel (NLP) de dernière génération. Il intègre des algorithmes sophistiqués qui lui permettent de comprendre et de générer du langage humain avec une précision et une fluidité remarquables. Cette capacité à interagir en langage naturel rend ChatGPT particulièrement performant pour une variété d'applications, allant de l'assistance clientèle à l'éducation.

La polyvalence de ChatGPT est un autre aspect de son avancée technique. Il peut être utilisé pour diverses tâches, comme la génération de texte créatif, la réponse à des questions, ou la rédaction de résumés. Cette flexibilité le rend applicable dans de nombreux domaines, ouvrant la voie à des utilisations innovantes.

ChatGPT se distingue également par sa capacité à mener des conversations interactives et à comprendre le contexte d'une discussion. Cette compréhension contextuelle lui permet de fournir des réponses plus pertinentes et personnalisées, améliorant ainsi l'expérience utilisateur.

### **L'ordinateur quantique**

L'ordinateur quantique représente une révolution technologique, promettant d'augmenter considérablement la puissance et l'efficacité du traitement informatique.



ARIANE  
ÉDITIONS