

## S'il te plait, ChatGPT, dessine-moi un mouton – 3<sup>ème</sup> article

### Synthèse des grandes questions que pose l'I.A.

*Si, adossé à son avion en panne au fond du désert, Antoine de Saint-Exupéry découvrait un dessin de ChatGPT, il serait sans doute bien dubitatif. De son côté, du haut de son étoile, le Petit Prince serait tout aussi interrogatif face à ce mouton issu d'une intelligence artificielle. Ce n'est pas le cas de Bill Gates qui voit en ChatGPT « la plus grande révolution industrielle depuis 1980 »<sup>1</sup>.*

*Passionné par ce sujet, j'essaye de compiler et d'écouter un maximum de ce qui se produit depuis que le tsunami ChatGPT secoue le monde entier. La compilation de dizaines d'articles de la presse quotidienne, du Figaro au Monde en passant par Les Échos ou La Croix, de dossiers spéciaux de tous les grands hebdomadaires et de nombreux podcast de journalistes pour une fois véritablement curieux et interrogatifs, n'est pas une démarche 100 % scientifique mais elle reflète, je l'espère, l'étendue des questions qui se posent et identifie les grandes thématiques qui sont en train de s'imposer.*

*Dans le premier article nous avons posé quelques bases de compréhension indispensables sur le fonctionnement de l'intelligence artificielle (IA) avant de présenter dans le second un état des réflexions concernant l'impact des IA sur le travail et l'emploi. Je propose dans ce 3<sup>ème</sup> article, une synthèse des questions qui se posent autour de l'I.A.*

Pour écrire cet article, j'ai utilisé l'I.A. J'ai utilisé un système de dictée, ce qui me permet d'aller plus vite dans la rédaction du premier jet. J'ai également demandé à ChatGPT (abonnement premium à ChatGPT4) de me proposer ses propres réponses à la question suivante : quelles sont les questions qui se posent concernant l'intelligence artificielle ? Pour les curieux, vous trouverez en annexe le dialogue que j'ai eu avec lui.

#### Fiabilité des données

Comme nous l'avons vu dans le premier article, l'IA n'est pas un système magique. C'est une construction de l'homme, de mathématiciens utilisant les statistiques et les probabilités, et d'ingénieurs poussant au maximum les capacités de calcul actuelles des outils informatiques. Ces systèmes apprennent, ils apprennent avec ce qu'on leur donne. C'est là que se pose la toute première question et certainement la plus importante : quelle est la base d'apprentissage ?

ChatGPT est, en partie, transparent sur le sujet. Sa base de connaissance est construite sur de l'information présente jusqu'à fin 2021 (et à partir de 2012). Il décline par exemple toute question sur la guerre en Ukraine. Il est en revanche plus compliqué de se faire un jugement sur la « pensée » qu'on lui a construite sur des aspects sociétaux ou politiques. L'avis semble assez largement partagé qu'il pense « américain », « californien » et un brin « woke ». A la limite, peu importe ce que pense ChatGPT. Comme tout le monde, une I.A. peut-être « woke »,

---

<sup>1</sup> Précision : il a investi 10 milliards dans la société Open AI qui a créé ChatGPT – Les Échos du 22 mars 2023

« intégriste musulman », « communiste » ou « libertarien », ce qui est important pour son utilisateur, c'est de le savoir. Et le problème est bien là. Qui aujourd'hui a la capacité de savoir ? De savoir faire le tri ? De comprendre que l'I.A. de Xi Ping, que celle de Google, que celle de Disney, que celle de l'OPEP ou de la FIFA ne pense(ront) que comme leurs géniteurs avec les arrière-pensées qui leur sont propres ?

Il est donc conseillé de bien garder en tête que quelle que soit la puissance de l'I.A., elle ne sait que ce que l'on a bien voulu lui apprendre.

## **Éthique et I.A.**

A partir du moment où l'on comprend qu'une I.A. générative apprend à partir de données qu'on lui donne, donc que l'on peut choisir, que ces données passent au travers des algorithmes que l'on peut biaiser, deux réflexes sont indispensables face aux résultats obtenus : être prudent et être critique. Si l'on prend un peu de recul, ce n'est pas pire, pas mieux, qu'avant ces I.A. : la Pravda n'était pas de droite, FoxNews n'est pas communiste et l'Équipe assez peu spécialiste de cosmologie. L'essentiel est de le savoir. Il est illusoire, utopique, de penser qu'une IA universelle et neutre pourrait exister ; si la question est de savoir si une I.A. peut être éthique, cela renvoie à une question de philosophie : c'est quoi l'éthique ?

Sur le sujet, Chat GPT nous dit de façon assez neutre que « si un modèle de langage est formé sur des données biaisées, il peut produire des résultats qui reflètent ces biais. Il est donc crucial de développer des méthodes pour détecter et corriger ces biais dans les systèmes d'I.A. ». Un système de données sera toujours biaisé et seul le roi des utopistes peut croire en « des méthodes pour détecter et corriger ces biais ». C'est le genre d'utopie qui a pu nous faire espérer que l'ONU serait éternellement un outil garantissant la paix dans le monde.

A un moment où les affrontements civilisationnels de tout ordre reprennent les armes, il est donc conseillé de retenir le principe qu'une I.A. a l'éthique de ses concepteurs.

## **Sécurité des données et I.A.**

Dès les premiers mois d'existence de ChatGPT3 et du tsunami qu'il a généré en début d'année 2023, il est clair qu'il n'y a pas de sécurité des données. Le cas Samsung a alerté très vite sur le sujet : des ingénieurs ont demandé à ChatGPT de les aider à résoudre un problème technique ; pour cela, ils ont donné au système des informations confidentielles, informations qui sont très vite ressorties ailleurs. Immédiatement, beaucoup de sociétés ont bloqué l'accès à ChatGPT.

La question qui se pose est donc de savoir comment utiliser les I.A. disponibles et garantir la sécurité de ses données. Nous avons vu que certains veulent un moratoire sur le développement de l'I.A. et que l'Europe veut produire une régulation de l'I.A. d'ici la fin de l'année ; nous ne sommes qu'au début des questions sur le sujet et très loin des réponses.

Début mai 2023, OpenAI, concepteur de ChatGPT, communiquait que son robot avait cessé de garder en mémoire les informations fournies par des utilisateurs professionnels. Mais attention,

par utilisateurs professionnels, il faut comprendre « ceux qui achètent la technologie OpenAI pour l'intégrer dans leurs propres solutions », pas des utilisateurs comme vous ou moi.

Il est conseillé d'être plus que prudent avec les informations que l'on fournit aujourd'hui aux I.A. ; changer des noms, arrondir des chiffres ne suffira pas à les tromper, artificielle certes, mais pas totalement bêtes non plus les I.A. !

### **Responsabilité de l'I.A.**

La question est la suivante : une I.A. vous a donné un conseil que vous avez suivi ; il s'avère que ce conseil était mauvais et vous a conduit à, sinon une catastrophe, au moins à un incident pour lequel vous êtes recherchés en responsabilité. Qui est responsable de l'incident ? Vous ? L'entreprise qui a conçu l'I.A. ? Celle qui a préparé la base d'apprentissage ? Celui qui a conçu les algorithmes ? voire l'ordinateur quantique qui commet encore quelques erreurs <sup>2</sup>? A ce jour, seules ces questions existent et il va falloir du temps avant qu'un début de réponse émerge.

Pour ceux qui sont dans la vie réelle de l'entreprise, il est pragmatique de partir du principe que vous restez 100 % responsable de vos décisions quelle que soit la méthode que vous utilisez pour réfléchir et décider.

### **Droit d'auteur et I.A.**

Ce point nous éloigne un peu des sujets qui intéressent le dirigeant d'entreprise. Je vous renvoie directement à la réponse de ChatGPT jointe en annexe. Sa réponse sur la question de savoir si l'I.A. peut être considérée comme créative est intéressante à découvrir.

### **Emploi et I.A.**

Vous avez certainement entendu parler de cette étude de la banque Goldman Sachs dont la presse relayait le chiffre de 300 millions d'emplois supprimés par l'I.A. Je vous renvoie sur ce thème à l'article précédent de cette série : <https://www.cogito-conseil.fr/chatgpt-ia-et-emploi-meme-pas-peur/>

### **Peurs et I.A.**

En mars, 1.300 personnalités du monde signaient une pétition pour demander un moratoire de six mois dans le développement des I.A. génératives. Depuis la Californie, ils venaient de réinventer le concept de l'auberge espagnole. Chaque signataire avait une raison différente de

---

<sup>2</sup> A ma connaissance, ChatGPT ne tourne pas sur des ordinateurs quantiques. Pas encore. Ces derniers n'ont pas encore la fiabilité nécessaire pour être utilisés « en production ».

demander cette pause. Mais qui peut croire que la machine peut s'arrêter ? Et qui peut penser que la Chine ou l'Inde vont attendre l'occident ?

Un des pères fondateurs de l'I.A., Geoffrey Hinton, lauréat du prix Turing en 2019 (le Nobel des informaticiens), annonçait en avril qu'il quittait Google. Cette décision était là pour « pouvoir parler des dangers de l'I.A. ». Curieusement, il n'est cependant pas signataire du moratoire. Nos chercheurs redécouvrent aussi le concept de l'apprenti sorcier.

De quoi ont peur ces personnalités au cœur de l'accélération de ces I.A. ? A lire plus précisément le moratoire ou leurs déclarations et en oubliant les arrière-pensées de guerre économique qu'ils se livrent dans l'ombre, on peut résumer ces peurs avec trois verbatims extraits de l'interview de G. Hinton au News York Times :

- « Il y a de profonds risques pour la société et l'humanité ».
- « Il est difficile de voir comment éviter que des mauvais acteurs l'utilisent pour de mauvaises choses » ».
- « L'idée que ce matériel puisse devenir plus intelligent que les gens n'est plus une utopie ».

G. Hinton et ses 1.300 amis me font penser à Robert Oppenheimer, le père de la bombe atomique qui après avoir conçu celles qui ont été larguées sur Hiroshima et Nagasaki en août 1945 s'est opposé au développement de l'arme nucléaire.

SI l'I.A. n'est pas une révolution, mais une évolution (voir mon 1<sup>er</sup> article sur le sujet), est-ce que nous sommes dans le même type de moment ? C'est ce que l'on peut penser de la réaction de ces personnalités. Comme le nucléaire qui a sa bombe mais qui nous produit de l'électricité, l'I.A. a deux faces, l'une est noire et inquiétante, l'autre est plus lumineuse et nous fera progresser.

## Conclusion

Au-delà de ces questions, de ces risques et de ces peurs dont on peut simplement espérer, à notre niveau, que les dirigeants et grands responsables du monde apportent une réponse, l'I.A. est entrée par la grande porte dans nos vies. Au sein de l'entreprise, il y a de la création de valeur à aller chercher avec ces outils. Nous verrons dans le prochain article comment apprendre à travailler avec ChatGPT & co.

Michel MATHIEU – Mai 2023

Tous les articles sont à retrouver sur <https://www.cogito-conseil.fr/blog/>

## Annexe : copie du dialogue avec ChatGPT 4 – 10 mai 2023

Peux-tu me rédiger un article de 1500 mots sur toutes les questions que pose le développement d'une IA comme ChatGPT ?

ChatGPT

Bien sûr, voici un article qui explore les nombreuses questions que pose le développement de l'IA, comme ChatGPT.

### **Titre: Explorer le labyrinthe: Les questions complexes posées par le développement de l'IA comme ChatGPT**

L'intelligence artificielle (IA) a fait des progrès remarquables ces dernières années, avec des modèles de langage comme ChatGPT de OpenAI qui peuvent générer des textes étonnamment cohérents et engageants. Cependant, ce développement rapide soulève de nombreuses questions complexes, allant des implications éthiques à l'impact sur l'emploi et à la sécurité des données.

#### **Éthique de l'IA**

L'une des questions les plus pressantes concerne l'éthique de l'IA. Comment pouvons-nous garantir que l'IA est utilisée de manière éthique et responsable ? De plus, comment pouvons-nous nous assurer que l'IA ne renforce pas les biais existants ? Par exemple, si un modèle de langage est formé sur des données biaisées, il peut produire des résultats qui reflètent ces biais. Il est donc crucial de développer des méthodes pour détecter et corriger ces biais dans les systèmes d'IA.

#### **Impact sur l'emploi**

L'IA a le potentiel de transformer le monde du travail. Certains emplois pourraient être automatisés, tandis que de nouveaux emplois pourraient être créés. Cela soulève des questions sur la façon dont nous devrions préparer les travailleurs pour ce changement. Quelles compétences seront les plus précieuses à l'ère de l'IA ? Comment pouvons-nous aider les travailleurs à faire la transition vers de nouveaux rôles ?

#### **Sécurité des données**

La sécurité des données est une autre préoccupation majeure. Les modèles de langage comme ChatGPT nécessitent de grandes quantités de données pour l'entraînement, ce qui soulève des questions sur la façon dont ces données sont collectées, stockées et utilisées. Il est essentiel de mettre en place des protections robustes pour garantir la confidentialité et la sécurité des données.

#### **Responsabilité**

Si une IA commet une erreur, qui est responsable ? Est-ce l'entreprise qui a développé l'IA, l'utilisateur qui l'a utilisée, ou l'IA elle-même ? Ces questions de responsabilité sont complexes et nécessitent une réflexion approfondie.

### **Réglementation**

Comment l'IA devrait-elle être réglementée ? Une réglementation trop stricte pourrait freiner l'innovation, tandis qu'une réglementation trop laxiste pourrait permettre des abus. Trouver le juste équilibre est un défi majeur.

### **Inégalités**

L'IA a le potentiel d'exacerber les inégalités existantes. Par exemple, si l'accès à l'IA est concentré entre les mains de quelques entreprises ou pays, cela pourrait créer des disparités significatives en termes de pouvoir et de richesse. Comment pouvons-nous garantir un accès équitable à l'IA ?

### **Conclusion**

Le développement de l'IA, comme ChatGPT, soulève de nombreuses questions complexes. Pour naviguer dans ce labyrinthe, nous aurons besoin d'une approche multidisciplinaire qui combine l'expertise en technologie,

Que dire de la question de l'IA et du droit d'auteur ?

### **ChatGPT**

La question de l'IA et du droit d'auteur est un domaine complexe et en constante évolution. Elle soulève plusieurs questions clés.

### **Création de l'IA**

L'une des principales questions concerne la création par l'IA. Si une IA, comme ChatGPT, crée une œuvre, à qui appartient le droit d'auteur ? Actuellement, la loi sur le droit d'auteur dans de nombreux pays stipule qu'une œuvre doit être créée par un être humain pour être protégée. Cela signifie que les œuvres créées par une IA ne seraient pas protégées par le droit d'auteur.

Pendant, cela soulève des questions sur le rôle de l'humain dans la création de l'IA. Si une personne a programmé l'IA ou fourni les données d'entraînement, ont-ils un droit sur l'œuvre créée par l'IA ? Et qu'en est-il de l'entreprise qui a développé l'IA ? Ces questions n'ont pas encore de réponses claires et pourraient nécessiter de nouvelles législations ou interprétations juridiques.

### **Utilisation des données d'entraînement**

Une autre question clé concerne l'utilisation des données d'entraînement. Si une IA est formée sur des données protégées par le droit d'auteur, cela constitue-t-il une violation du droit d'auteur ?

Actuellement, la réponse à cette question varie selon les juridictions et dépend souvent de facteurs tels que la quantité de données utilisées et le but de leur utilisation.

### Conclusion

La question de l'IA et du droit d'auteur est complexe et sans précédent. Alors que l'IA continue d'évoluer, il sera crucial de continuer à explorer ces questions et à développer des lois et des réglementations qui protègent à la fois les droits des créateurs et encouragent l'innovation.

Peut-on considérer qu'une IA est créative à partir du moment où elle ne fait que refaire ce qu'elle a appris ?

### ChatGPT

La question de savoir si une IA peut être considérée comme créative est un sujet de débat parmi les experts en IA, en philosophie de l'esprit et en psychologie cognitive. Voici quelques perspectives à considérer.

1. **L'IA comme outil:** Dans cette perspective, l'IA est considérée comme un outil avancé qui aide les humains à accomplir des tâches plus efficacement ou à générer des idées que les humains peuvent ensuite explorer. Dans ce cadre, la "créativité" de l'IA est en réalité le produit de la programmation humaine et de l'entrée de données. L'IA, comme ChatGPT, génère du texte en se basant sur les patterns qu'elle a appris de ses données d'entraînement, mais elle ne comprend pas ce qu'elle écrit et ne peut pas avoir d'intentions créatives au sens où nous l'entendons pour les humains.
2. **L'IA comme co-créatrice:** Dans cette perspective, l'IA est vue comme un partenaire créatif qui peut générer des idées nouvelles et utiles. Cette vision reconnaît que bien que l'IA apprenne à partir des données d'entraînement fournies par les humains, la manière dont elle combine ces informations peut conduire à des résultats inattendus et novateurs.
3. **L'IA comme entité créative:** Cette perspective est plus radicale et suggère que l'IA pourrait être considérée comme une entité créative à part entière. Cependant, cela soulève des questions philosophiques et éthiques complexes, car la créativité est généralement associée à des aspects de l'expérience humaine tels que la conscience, l'intentionnalité et l'émotion, que l'IA ne possède pas actuellement.

Il est important de noter que la majorité des chercheurs et des experts en IA s'alignent sur les deux premières perspectives. Les algorithmes d'IA, même les plus avancés, sont fondamentalement des outils qui opèrent selon les instructions programmées par des humains et qui apprennent à partir des données fournies par des humains. Bien qu'ils puissent générer des résultats surprenants, novateurs ou semblent "créatifs", cela ne signifie pas nécessairement qu'ils possèdent la créativité au sens humain du terme.