

Que penser de l'Intelligence Artificielle ?

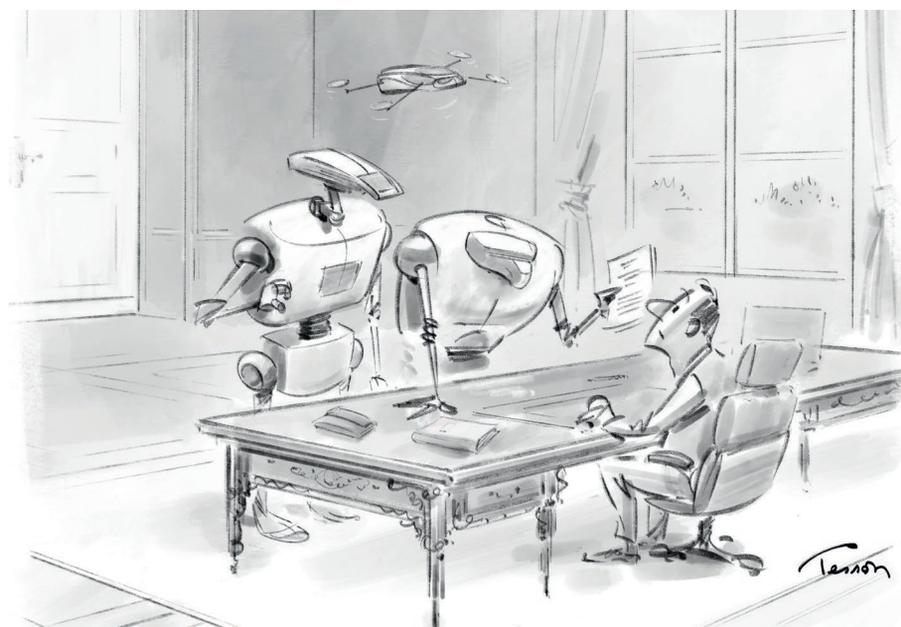
QU'EN DIT-ON ?

“ L'Intelligence Artificielle, quel exploit !”

“ Les robots nous feront la peau.”

“ L'IA, c'est pour nous manipuler.”

“ Rien ne remplacera l'intelligence humaine.”



— On vous a rédigé un projet pour la loi de programmation militaire : on a besoin de dix nouveaux bataillons de robots et d'une grande flotte de drones.

L'ÉDITO

Le développement spectaculaire de l'Intelligence

Artificielle (IA) est porteur

de bienfaits potentiels

considérables. Mais elle tend

à inquiéter par sa puissance

et son opacité, et par la remise

en cause possible de la place

de l'intelligence humaine et,

par là, de l'homme lui-même.

Comment comprendre le

phénomène de l'IA et le

maîtriser au mieux au service

de l'homme ?

LE CONSEIL SCIENTIFIQUE

Qu'est-ce que l'IA ? Peut-on la maîtriser ?

UNE INTELLIGENCE, VRAIMENT ?

Le terme intelligence est compris de multiples façons. Selon une optique matérialiste, répandue dans les milieux traitant de l'IA, l'intelligence est la capacité à résoudre des problèmes et à progresser dans cette solution. Dans cette perspective, la capacité de traitement de machines n'est pas en soi différente de celle de l'homme, et l'intelligence humaine peut être dépassée. Ce serait notamment le cas si l'IA atteignait le niveau appelé 'superintelligence', où elle pourrait apprendre à construire elle-même des IA encore supérieures, par un processus cumulatif. Une telle perspective peut donner le vertige.

Mais une saine philosophie, notamment nourrie par la foi chrétienne, ne peut s'arrêter à cette conception de l'homme. Ne réduisant pas l'esprit et l'intelligence à la solution technique de problèmes, elle pourra envisager que des machines aient une capacité de traitement très supérieure à celle de l'homme sans conclure à l'infériorité de ce dernier. L'esprit humain possède en effet une supériorité ontologique radicale par sa capacité à penser consciemment, ainsi que par sa liberté et sa créativité. Par ailleurs, l'homme est un être relationnel. En bref, c'est une personne. Par là, lui seul est ouvert à l'illimité et notamment à l'accueil de la transcendance divine.

DES MACHINES, MAIS SURPUISSANTES ?

Avec l'IA, une machine traite notamment des données immenses à l'aide d'algorithmes basés sur des calculs de probabilité. Sa capacité est sans précédent, non seulement par sa puissance de calcul, mais aussi par ses techniques d'apprentissage qui multiplient son efficacité. Cela explique qu'elle puisse battre des humains aux échecs ou au go, et cela par des moyens non prévus à l'avance par les concepteurs. Cela frappe le public et on admire cette 'intelligence'. Mais, ce faisant, la machine ne déploie en rien une pensée, moins encore une capacité à déceler le sens possible des choses et leur signification morale en termes de bien et de mal, et par là à agir dans la liberté en engageant sa responsabilité. En réalité, elle explore un univers de possibilités immense, mais sans création ou imagination au-delà d'une prodigieuse combinaison

« L'esprit humain possède une supériorité ontologique radicale par sa capacité à penser consciemment, ainsi que par sa liberté et sa créativité. »

de données. Que l'homme soit battu aux échecs prouve simplement que dans ce jeu la puissance de calcul l'emporte sur la pensée et l'imagination. De ce point de vue, les craintes de ceux qui voient émerger une superintelligence dominant l'homme sont infondées. Cela dit, cette crainte garde tout son sens au vu de l'extraordinaire puissance de ces outils. Au niveau ultime, on peut concevoir une IA intégrée surpuissante qui exercerait une domination catastrophique sur l'humanité. Non comme une intelligence consciente, mais comme une machine échappant au contrôle de ses concepteurs et poursuivant les instructions inscrites dans son programme au-delà de toute rationalité. Ce serait à la façon de l'apprenti sorcier du poème de Goethe, où un balai mobilisé magiquement pour aller chercher

de l'eau finit par se surmultiplier et inonder la maison. Il reste que les services que l'IA peut rendre sont immenses et que son développement rapide est inéluctable. Sans qu'on doive ici les développer, d'ores et déjà chacun les utilise chaque jour et s'en félicite : moteur de recherche, recherche d'itinéraire, traduction, rédaction, diagnostic de santé, etc.

LES PRINCIPAUX ENJEUX ÉTHIQUES DE L'IA

Le questionnement éthique à propos de l'IA porte sur trois niveaux : les données, les algorithmes et la responsabilité.

Les données d'abord sont sa matière première. D'où une série de questions : comment sont-elles collectées ? Quel est le degré de contrôle des personnes sur cette collecte et l'usage qui en est fait ? Mais aussi : quelle est la fiabilité de ces données, leur vérification, quels biais implicites ? Des IA utilisant les données issues de réseaux sociaux se sont retrouvées émettre des conclusions racistes.

Les algorithmes ensuite sont au cœur de l'IA. Entre autres, un risque majeur est ici le manque de transparence à l'égard des utilisateurs, comme on le voit avec les algorithmes des réseaux sociaux, qui livrent à leur public un flux d'images et de textes selon une logique opaque pour eux. Mais un plus grand risque est l'émergence de conséquences nullement prévues par qui que ce soit. Un outil d'octroi de crédit bancaire peut avoir des effets d'exclusion raciale. L'algorithme

ne verra pas ce qu'il fait, car à nouveau il ne pense pas. La responsabilité enfin est une dimension majeure puisque les machines n'en sont pas elles-mêmes dotées. Il est donc vital que la responsabilité soit clairement identifiée là où elle se situe, et donc au niveau humain.

S'y ajoutent des enjeux de stabilité des systèmes et plateformes et des enjeux éthiques environnementaux liés à l'énergie déployée pour maintenir et faire fonctionner les IA.

DES EFFETS COLLECTIFS DE NATURE NOUVELLE ET POTENTIELLEMENT DANGEREUX

Ceci vaut déjà au niveau des acteurs : personnes, entreprises et organisations. Mais une telle réflexion est d'autant plus urgente qu'on voit poindre des effets collectifs potentiellement nocifs - à commencer par l'utilisation militaire. On ne peut qu'évoquer rapidement ces points.

On a parlé des réseaux sociaux : leur rôle culturel est désormais majeur, mais l'usage qui en est fait n'est pas globalement positif, notamment par les jeunes (Tiktok, Instagram, etc.). Ces algorithmes tendent à replier chacun sur un modèle statistique moyen et par là uniformisant, soit en général, soit niche par niche, parquant à la limite les utilisateurs dans de petits mondes consensuels et mutuellement hermétiques. On peut aussi citer les capacités nouvelles de circulation d'informations tendancieuses ou manipulées.

A plus large échelle, on peut évoquer une capacité considérablement supérieure de contrôle social, comme en Chine, avec identification faciale et caméras omniprésentes, localisation potentiellement permanente, notation sociale régulant l'accès aux biens et services, etc. Or cela ne résulte pas du seul totalitarisme du régime : ses motivations et ses promesses rencontrent celles faites ailleurs (services rendus et sécurité).

Il faut enfin évoquer la question de l'emploi. A l'instar de la robotisation, l'IA peut avoir un impact important, mais cette fois ce sont des emplois qualifiés qui seraient touchés. Un risque possible serait qu'une minorité s'en tire bien, mais qu'une majorité soit déclassée voire exclue.

LES CONDITIONS D'UN BON USAGE

La question concrète est celle de savoir comment utiliser les bienfaits de l'IA sans souffrir de ses risques, et comment la mettre au service du développement humain intégral. Cela mobilise, comme toujours, le double niveau de l'éthique des personnes et des organisations, et de la réglementation.

Cela vise bien sûr d'abord les principes généraux devant présider à toute action pouvant affecter d'autres personnes ou la société : respect de la personne, recherche de son développement plénier, attention particulière aux plus vulnérables et au risque de dépendance, absence de discrimination, et plus généralement tous les principes de responsabilité sociale et environnementale.

Mais, outre des qualités particulières de robustesse technique et de fiabilité, un point essentiel et spécifique à l'IA est le maintien d'un contrôle responsable de son fonctionnement. Cela vise d'abord la transparence et la traçabilité de la collecte des données, et notamment

le respect de la vie privée, donnant aux personnes la maîtrise de leurs données personnelles et la garantie de leur utilisation responsable. Cela vise ensuite le fonctionnement des algorithmes, tant *ex ante*, dans le dessein du système, qu'*ex post* par des tests réguliers et des études d'impact de ce que

produit effectivement ce système, qui ne doit pas devenir opaque. Le tout avec responsabilisation des concepteurs comme des utilisateurs.

UN PROGRÈS POSSIBLE DANS LA COMPRÉHENSION DE L'HOMME

Au total, l'IA est un phénomène considérable, potentiellement source de grands bienfaits et de conséquences très négatives : en cela, on retrouve le questionnement séculaire sur l'innovation technique. Mais la question prend ici un tour nouveau, du fait de la puissance de l'IA et de sa capacité à bouleverser l'homme et la société. Parallèlement, outre l'apport des neurosciences, cela peut conduire aussi à une meilleure compréhension de la nature véritable de l'intelligence humaine et de l'esprit humain, créé, rappelons-le, à l'image de Dieu, et qui ne se réduit pas à une puissance de calcul. ●

À RETROUVER SUR WWW.PROPERSONA.FR

En bref

COMMENT COMPRENDRE LE PHÉNOMÈNE DE L'IA ET LE MAÎTRISER AU MIEUX AU SERVICE DE L'HOMME ?

C'est par une compréhension erronée de la nature de l'intelligence humaine qu'on conclut parfois à son infériorité par rapport à l'IA, du fait que celle-ci effectue des calculs complexes avec une bien plus grande efficacité. Mais si, contrairement à l'homme, l'IA ne pense pas à proprement parler, sa puissance et sa relative opacité exigent une réflexion et une action résolues sous responsabilité humaine pour en contrôler les effets.

À RETROUVER SUR WWW.PROPERSONA.FR



La citation

L'Intelligence Artificielle est [...] un outil sui generis. Alors que l'utilisation d'un outil simple (comme le couteau) est sous le contrôle de l'être humain qui l'utilise [...], l'Intelligence Artificielle, en revanche, peut s'adapter de manière autonome à la tâche qui lui est assignée. [...] Face aux prodiges des machines, qui semblent capables de choisir de manière autonome, nous devons être clairs sur le fait que la décision doit toujours être laissée à l'être humain. »

PAPE FRANÇOIS, DISCOURS À LA SESSION DU G7 SUR L'INTELLIGENCE ARTIFICIELLE, 13-15 JUIN 2024.

Pour aller plus loin

PAPE FRANÇOIS, Message pour la 58^e journée mondiale des communications sociales : Intelligence Artificielle et sagesse du cœur, 24 janvier 2024.

PAPE FRANÇOIS, Message pour la 57^e journée mondiale de la paix : Intelligence Artificielle et paix, 1^{er} janvier 2024.