

L'Homme face à l'intelligence artificielle : repenser l'éthique de la relation homme-machine

Par Paul-Oliver GIBERT

Président de Digital & Ethics

En cette deuxième décennie du XXI^e siècle, l'intelligence artificielle semble incarner l'angoisse prométhéenne liée au progrès technique et scientifique : « Une fois que les hommes auraient développé l'intelligence artificielle, celle-ci décollerait seule, et se redéfinirait de plus en plus vite... Les humains, limités par une lente évolution biologique, ne pourraient pas rivaliser et seraient dépassés », a déclaré, lors d'une interview, Stephen Hawking⁽¹⁾.

L'intelligence artificielle peut-elle permettre aux machines de dépasser les humains, leurs créateurs ? Certains le redoutent, d'autres, notamment les adeptes du trans-humanisme, espèrent que, combinée avec d'autres technologies, l'intelligence artificielle permettra une « sortie par le haut de la condition humaine ». Entre ces deux points extrêmes existe une relativement grande variété de positions.

Les débats sur les espoirs et craintes liés à l'intelligence artificielle occupent une très grande place dans les médias mais force est de constater que les applications pratiques et concrètes sont encore loin de permettre aux « agents intelligents » de rivaliser pleinement avec les êtres humains. S'il est vrai que des machines, des algorithmes mis en œuvre par ordinateur peuvent battre aux échecs, au jeu de go ou au poker des champions humains, il s'agit de domaines très spécifiques et limités de l'activité humaine.

Toutefois, nous constatons aujourd'hui une rupture dans les relations entre l'homme et la machine, rupture permise par l'intelligence artificielle : le recours au langage naturel, notamment sous sa forme parlée, pour échanger avec des machines qui ont une capacité d'apprentissage autonome met l'humain en position de déléguer certaines actions à ces machines autonomes.

Cette rupture appelle doublement une réflexion éthique : quel doit être, quel peut être le bon usage de cette nouvelle possibilité offerte par les machines ? Comment, de façon efficace, peut-on permettre la réalisation des espoirs de certains dans l'intelligence artificielle tout en évitant les cauchemars redoutés par d'autres ?

Le développement de l'intelligence artificielle permet de déléguer des décisions à des machines, ce qui renouvelle l'éthique de la relation homme-machine

L'intelligence artificielle est un ensemble très vaste et très hétérogène : la simulation de l'intelligence peut prendre différentes formes. Il peut s'agir d'une machine dont l'architecture reproduit la structure du cerveau comme dans le cas des réseaux de neurones⁽²⁾ afin d'obtenir de meilleures

(1) http://www.lemonde.fr/pixels/article/2014/12/03/hawking-l-intelligence-artificielle-pourrait-mettre-fin-a-l-hu-manit533135_4408996.html#TezizU9P6mRfruAf.99

(2) BLAYO F. et VERLEYSEN M., *Les Réseaux de neurones artificiels*, Presses Universitaires de France, Que sais-je n°30421, réédition 1996.

performances. Il peut aussi s'agir d'une simulation lors de l'interface entre la machine et l'homme : il s'agit notamment de la communication en langage naturel (de tous les jours), ou encore de la reconnaissance et de la synthèse vocale avec notamment SIRI et Google Now.

Toutefois, la simulation de l'intelligence artificielle est parfois elle-même simulée : « Concevoir un code qui a l'air intelligent n'a rien à voir avec une machine qui apprend toute seule. On lui fixe des règles et on programme la réponse que la machine doit avoir si elle rencontre ces différents cas de figure. C'est ce que l'on retrouve dans 90 % de ce qui est vendu aujourd'hui comme de l'intelligence artificielle », écrit Thomas Gouritin⁽³⁾.

Malgré ces arrangements, la rupture est réelle dans la relation entre l'homme et les machines : il devient possible de lui déléguer certaines actions dont on ne connaît pas *a priori* le résultat en lui parlant comme on peut le faire avec un être humain.

Les progrès réalisés dans le traitement automatisé du langage naturel permettent aux machines d'interpréter le langage que nous employons pour le traduire en instructions qu'elles peuvent traiter. Cette évolution perceptible désormais dans les assistants installés sur les *Smartphones* (Siri ou Google Now, notamment) a une force symbolique très grande. Le test de Turing, imaginé par Alan Turing en 1950 pour distinguer un raisonnement humain de celui d'un ordinateur, s'appuie sur la capacité de celui-ci à reproduire la sémantique d'une conversation humaine⁽⁴⁾.

D'année en année, la capacité de la machine à tenir un « dialogue » s'est développée, entretenant notamment l'idée d'une compétition entre l'homme et les ordinateurs.

Cette rupture n'est pas sans susciter des interrogations sur le bon usage de ces nouvelles techniques : « Des centaines d'ingénieurs et universitaires ont signé cette semaine une lettre ouverte pour alerter sur les dérives d'une intelligence artificielle (IA), dont les progrès techniques seraient réalisés en dehors de toute considération sur les progrès sociétaux qu'elle doit apporter⁽⁵⁾ ».

Disposant d'ordinateurs et de procédés de traitements plus puissants (ceux du *Big data* notamment), les techniques de l'intelligence artificielle permettent de développer des outils ayant une certaine capacité d'apprentissage et de fonctionnement autonome. Il s'agit là aussi d'une vraie rupture.

Très schématiquement, il est possible de considérer qu'il y a, à ce jour, quatre stades dans la relation entre l'homme et ses relais techniques :

- le premier, celui de « l'outil à main », qui ne produit de résultat qu'actionné directement par un humain ;
- la machine qui, utilisant une ressource extérieure (énergie et/ou capacité de calcul), réalise les tâches pour lesquelles elle est directement dirigée et qui peut la poursuivre par un effet d'inertie ;
- la machine qui, utilisant une ressource extérieure (énergie et/ou capacité de calcul), réalise les tâches pour lesquelles elle a été programmée et qui peut la poursuivre par un effet d'inertie ;
- la machine qui dispose de capacités d'acquisition d'information et d'apprentissage qui permettent de réaliser des tâches pour lesquelles elle n'a pas été explicitement programmée.

Ce dernier stade est celui qui est spécifiquement rendu possible par l'intelligence artificielle et il s'agit d'une profonde rupture dans la relation entre l'homme et la machine. La catégorie des ma-

(3) <https://www.frenchweb.fr/larnaque-chatbots-durer-a-t-elle-encore-longtemps/305697>

(4) TURING A., *Computing Machinery and Intelligence*, (en), Oxford University Press, vol. 59, n°236, octobre 1950, p. 433-460.

(5) <http://www.numerama.com/magazine/31868%2Dlettre%2Douverte%2Dsur%2DI%2Dintelligence%2Dartificielle%2Dia%2Det%2Dsa%2Dplace%2Ddans%2Dla%2Dsociete.html>

chines autonomes permet et va permettre de développer des relations par lesquelles on confie à l'une de ces machines de troisième génération des tâches intégrant une prise de décision.

De nombreux outils numériques permettent aujourd'hui, en recourant notamment à des outils d'intelligence artificielle, de déléguer certaines actions nécessitant une certaine forme de prise de décision.

Nous sommes très proches de ce qui existe entre humains, notamment dans le cadre d'une délégation de pouvoir. En droit, la délégation de pouvoir consiste pour une personne (par exemple, un chef d'entreprise) à confier l'exercice de pouvoirs dont il est titulaire à une autre personne (un de ses cadres) qui les exerce sous sa responsabilité. En d'autres termes, déléguer c'est donner à un autre le pouvoir de décider à sa place. Nous sommes à l'opposé de l'instruction donnée à un exécutant.

De façon courante et très visible, il s'agit notamment des calculateurs d'itinéraires et de guidage par GPS qui intègrent notamment l'état de circulation et qui permettent ainsi de minimiser le temps de trajet ou la distance à parcourir. Un peu comme pour le jeu de go, le calculateur d'itinéraire dépasse l'humain par sa capacité à intégrer en temps réel des informations sur les embouteillages, et, en particulier, celles qui sortent du champ de perception du conducteur.

Une prochaine étape de cette délégation serait celle du véhicule autonome qui prendrait en charge la totalité des décisions liées à l'exécution du déplacement, y compris les décisions lors de situations de crise, notamment lors d'accidents. Des initiatives ont été prises pour résoudre ce dilemme éthique, parmi lesquelles la « moralmachine ⁽⁶⁾ » qui compile les points de vue humains sur les décisions morales que devraient prendre les machines autonomes dans ces situations.

La démarche éthique est particulièrement adaptée pour rechercher le bon usage des délégations à des machines permis par les progrès récents des techniques d'intelligence artificielle

L'approche éthique s'avère particulièrement pertinente pour identifier aujourd'hui les règles qui devraient régir le comportement de ces machines autonomes : ces technologies sont récentes et leurs réalisations encore largement expérimentales, ce qui justifie une approche pragmatique.

Quelle est la spécificité de l'approche éthique par rapport aux approches juridiques d'une part, morales et déontologiques d'autre part ? Ces trois notions sont complémentaires (voir le schéma). Le droit et la déontologie reposent sur un ensemble de normes, de règles ou de lois dont le respect s'impose à tous (droit) ou à tous les membres d'une profession (déontologie). La morale repose sur des règles établies issues d'un système de valeurs qui peut être issu d'une religion (morale chrétienne, par exemple). L'éthique repose sur une réflexion, un questionnement sur la situation qu'elle examine. Cette démarche trouve son utilité lorsque les démarches juridiques et morales peinent à fournir des réponses à la question du bon usage d'un objet ou du bon comportement face à une situation.

Le recours à une approche éthique est ainsi particulièrement adapté pour répondre aux enjeux du bon usage des machines autonomes nées de l'intelligence artificielle, tant les cas d'utilisation sont divers. D'où la pratique professionnelle de l'auteur de ces lignes, dans le cadre du recours à ces technologies, de poser quatre types de questions éthiques :

- quelle est la mission confiée à la machine ?
- quelle est la fiabilité des résultats qu'elle produit ?
- quelle est l'information des individus ?
- quel est le consentement des individus ?

(6) <http://moralmachine.mit.edu/hl/fr>

Ethique, morale, droit et déontologie

Des notions complémentaires



© Digital Ethics - Présentation DEL

Pour illustrer de façon concrète ces questions, nous avons extrait d'un article certains domaines dans lesquels sont attendus les impacts des développements des techniques d'intelligence artificielle. Cela permet d'en saisir les enjeux économiques et sociaux⁽⁷⁾ et de mettre en regard les questions éthiques qu'elles appellent :

« - *L'intelligence artificielle remodèle complètement les sciences de la vie, la médecine et les soins de santé en tant qu'industrie. Les innovations dans l'intelligence artificielle font avancer l'avenir de la médecine de précision et de la gestion de la santé de la population de façon incroyable.* »

Quels sont les objectifs du recours à l'intelligence artificielle : soigner l'individu ou prendre en compte les impératifs de santé publique ? Dans ce deuxième cas, comment seraient traités des comportements individuels à risque ?

« - *Des chercheurs en Espagne et au Portugal ont utilisé des réseaux de neurones artificiels pour prévoir les fluctuations de prix et d'utilisation des réseaux électriques : faire des prédictions adéquates basées sur les modèles de consommation et de disponibilité permet d'atteindre une efficacité et des économies de coûts bien supérieures.* »

Quelle est la fiabilité des prédictions ? N'y a-t-il pas un risque de voir un système construit autour de prédictions auto-réalisatrices ? Comment se feront les arbitrages entre besoins d'électricité et capacités de production ? les personnes pourront-elles encore décider de la température qu'il fait dans leur résidence ?

« - *Les scientifiques de Harvard ont utilisé le deep learning pour effectuer des calculs viscoélastiques, ce sont les calculs utilisés pour les prédictions de tremblements de terre. Leur travail a amélioré le temps de calcul de 50 000 %. Quand il s'agit de calcul de tremblement de terre, le calendrier est important et cette amélioration peut être essentielle pour sauver des vies.* »

Là aussi se pose la question de la fiabilité des prévisions et de l'éventuelle automatisation de l'alerte : quel sera le risque d'être réveillé en pleine nuit par une fausse alerte ?

(7) <http://knowstartup.com/2017/08/10-amazing-ways-deep-learning-will-rule/>

« - Les entreprises qui conçoivent les voitures autonomes doivent enseigner à un ordinateur comment prendre en charge des situations critiques en utilisant des systèmes de capteurs numériques au lieu des sens humains. Ces nouveaux services pourraient fournir des modèles d'affaires inattendus aux entreprises. »

Nous retrouvons ici la question des situations extrêmes dont la charge éthique remonte de l'utilisateur au concepteur/programmeur du véhicule, mais aussi de la fiabilité de la machine : les capteurs numériques permettront-ils d'avoir une bonne perception de la situation ? Quels seront les principes sous-jacents du pilotage des véhicules : transporter des individus en les déchargeant de la conduite ou optimiser globalement la circulation au sein d'une agglomération ?

Il ressort à la fois la puissance des changements potentiellement portés par l'intelligence artificielle, mais aussi le fait que ceux-ci ne sont pas dissociables des changements portés par la révolution numérique⁽⁸⁾, si ce n'est sur un point : le développement d'outils auxquels il est possible de déléguer certains types de décisions.

Les machines autonomes réputées intelligentes doivent intégrer dans leur conception des règles, des contraintes éthiques. Ces règles sont nécessaires pour que les pouvoirs que les humains leur délégueront soient exercés conformément à ce que nous attendons. L'un des avantages de ces technologies est que ces engagements éthiques pourront être tracés et vérifiés.

Toutefois, ces mesures n'auront pas de plein effet si les humains renoncent à leur libre arbitre : la délégation d'un pouvoir n'est pas son abdication. Il faut, au contraire, renforcer et développer la capacité de tous les hommes à juger et à agir en pleine conscience et à ne pas se laisser asservir par des machines qu'ils auront eux-mêmes construites.

(8) COLIN N. et VERDIER H., *L'Âge de la multitude. Entreprendre et gouverner après la révolution numérique.*