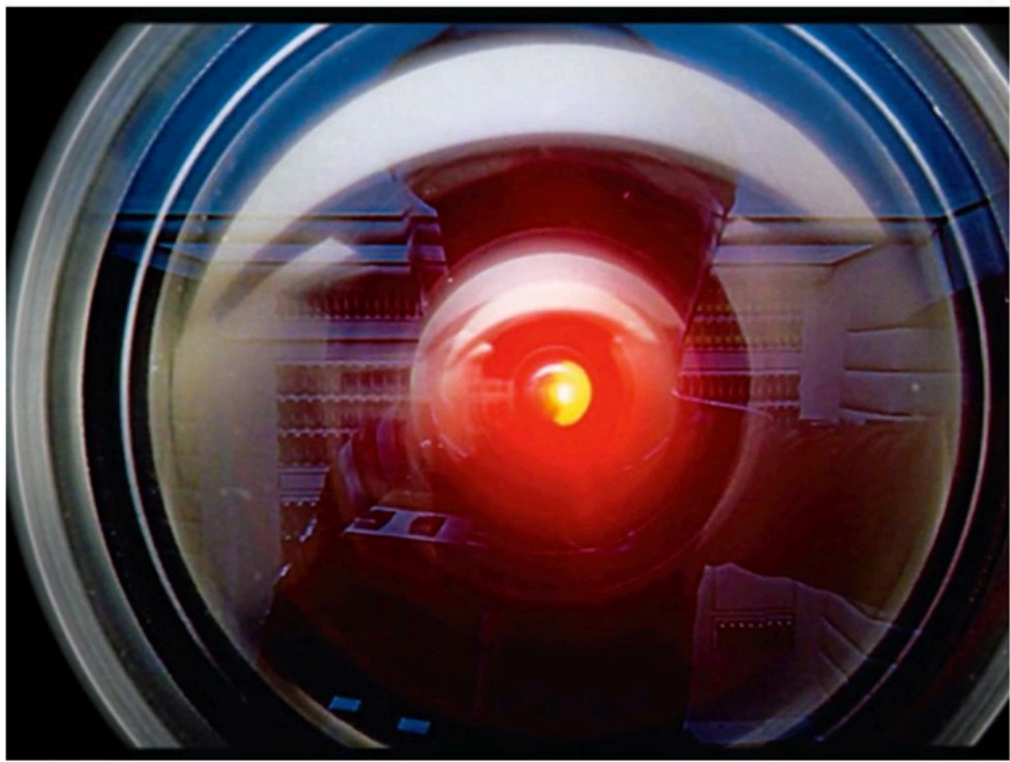


Intelligence artificielle: Darwin ou Kubrick ?

Par Challenges.fr le 15.06.2019 à 15h27

ABONNÉS

TRIBUNE - L'IA est-elle l'inéluctable aboutissement d'un processus darwinien commencé par la sélection biologique, suivie de la sélection des meilleures sociétés et aujourd'hui des meilleurs algorithmes ? Eclairage par le philosophe Guillaume Von der Weid, spécialiste des questions IA et santé.



Capteur visuel de HAL 9000, l'ordinateur du vaisseau Discovery One dans 2001 : L'odyssée de l'espace. Le film de Stanley Kubrick, sorti en 1968, interrogeait déjà les limites de l'intelligence artificielle.

MGM

SUR LE MÊME SUJET

• **Et si l'intelligence artificielle n'existait pas ?**

L'intelligence artificielle est devenue une source exponentielle du progrès technologique. Commerce, sécurité, transport, éducation, santé divertissement, ses applications semblent infinies. Avec le même optimisme dont font preuve les transhumanistes quand ils nous promettent "la mort de la mort", les programmeurs nous annoncent la fin de l'ignorance et l'optimisation du monde. Serions-nous arrivés à la fin de l'Histoire, qui ne serait plus l'objet de grands récits et de révolutions sanglantes mais seulement d'analyses de mégadonnées et d'améliorations incrémentales ? L'intelligence, en s'affranchissant de nos reliefs cérébraux, parviendrait-elle à écrêter les contradictions du progrès ?

Une ou des réalités ?

L'intelligence artificielle (IA) peut en effet être envisagée comme le prolongement du processus darwinien qui aurait commencé par la sélection biologique des meilleurs gènes, se serait poursuivi par la sélection culturelle des meilleures sociétés et se propagerait désormais par la sélection technologique des meilleurs algorithmes. L'évolution naturelle toucherait ainsi à sa troisième et dernière étape, dépassant à la fois la thèse biologique et l'antithèse culturelle dans un artefact électronique délivré des blocages de l'ignorance et de la superstition. L'IA, par une plasticité lui permettant de s'appliquer à toutes les dimensions de l'action humaine, en compilant le réel tout entier, notamment par la production de "jumeaux numériques", constituerait l'ultime stade de la rationalité.

Encore faut-il s'entendre sur ce qu'on appelle le réel. Car il est loin de faire l'unanimité ! La réalité suprême de la théologie n'est plus aujourd'hui qu'une abstraction vieillotte. L'obésité qui était considérée comme un signe de richesse est aujourd'hui traitée comme une maladie, l'homosexualité, vice inexpiable, est devenue une condition naturelle. Le travail était un signe d'esclavage, il est désormais le seul moyen légitime de s'élever socialement. La Terre était au centre d'un univers fini, l'espace et le temps étaient des absolus, l'humain était au sommet de la nature. Thomas Khun montre ainsi comment la connaissance du monde progresse de deux façons très différentes, soit en mode "normal", comme précision et approfondissement d'une vision du monde donnée, ainsi les recherches scientifiques à l'intérieur du "paradigme newtonien", soit en mode "révolutionnaire" lorsqu'une découverte bouleverse, non pas les résultats antérieurs, mais le cadre même des calculs. On est ainsi passé d'un monde clos à un univers infini avec Galilée, d'un espace orthonormé à un espace courbe avec Einstein, d'une théorie fixiste des espèces à une théorie de l'évolution avec Darwin. L'IA peut certes pousser à la perfection la science normale, mais non provoquer des révolutions scientifiques qui impliquent de changer le cadre de la représentation, c'est-à-dire ce qu'on entend par réalité. Et cette limite de l'univocité s'applique encore davantage aux valeurs humaines où le normal n'est plus un état descriptif de corrélations, mais un jugement prescriptif de comportements. L'abolition de la peine de mort, par exemple, est un changement de référentiel moral et non l'optimisation d'un système judiciaire donné.

Jusqu'où optimiser ?

Kubrick a résumé cette opposition dans 2001 L'Odyssée de l'espace, où des singes préhistoriques, confrontés à l'apparition d'un monolithe mystérieux, aux formes parfaites, vont en tirer l'invention de la technique par une sorte de projection de cette perfection sur les objets qui les entourent, soudain perçus comme des instruments polyvalents, ainsi un os qu'un des singes va utiliser comme massue. Par une ellipse audacieuse, l'os se transforme en satellite, et l'intelligence du singe qui a touché le monolithe en intelligence artificielle, intégrée à une mission spatiale. Or l'IA du vaisseau, nommée CARL, jugeant le facteur humain trop aléatoire pour une mission d'une telle importance, tente d'en éliminer les humains.

Deux idées émergent du récit. L'ellipse montre que l'intelligence instrumentale, aussitôt apparue, est inéluctablement conduite à se reproduire à l'extérieur d'elle-même sous forme d'IA. La décision inhumaine de CARL est, quant à elle, l'aboutissement logique de cette recherche d'efficacité à tout prix, au détriment des limites et des valeurs qui rattachent littéralement les moyens qu'on emploie aux buts qu'on poursuit. À quoi bon mener une mission spatiale qui élimine ceux pour qui elle est réalisée ? À quoi bon collecter des données de santé si cela rend la société malade ? Créer des robots tueurs si c'est pour être moins en sécurité ? Le propre de l'intelligence est d'améliorer nos buts plus que nos moyens, précisément parce que les buts déterminent les moyens et non l'inverse. C'est pourquoi l'enfantement, qui n'est en un sens qu'un moyen de perpétuer l'espèce, est un événement magnifique : il donne naissance à une nouvelle vision du monde, unique, et pas seulement à des données (à exploiter rationnellement) et des besoins (à satisfaire efficacement).

L'intelligence artificielle doit donc toujours être pensée, non comme un but mais comme un instrument, avec ses dangers, son coût humain et social, sa vision implicite du monde. Elle doit obéir à nos buts et non, à l'inverse, les englober dans le monde meilleur où elle finira par nous faire disparaître.

Par Guillaume Von der Weid, philosophe