



LES TROIS VISAGES DE L'INTELLIGENCE ARTIFICIELLE

LUC FERRY
ÉCRIVAIN ET PHILOSOPHE

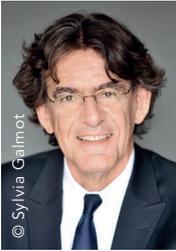
Le sujet devient si important, les progrès de l'intelligence artificielle (IA) sont si impressionnants, si bouleversants au sens propre du terme, qu'il nous faut y revenir et suivre ses avancées si nous voulons bien comprendre à quoi nous avons affaire. Car il existe trois sortes d'IA, trois visages différents

**« LA MACHINE APPREND TOUTE SEULE,
EN PERMANENCE, SANS JAMAIS
S'ARRÊTER DE PROGRESSER. »**

de cette nouvelle forme de calcul, sinon de pensée, et il est tout à fait essentiel de ne pas les confondre si l'on veut comprendre la révolution qui est d'ores et déjà en marche. Il y a d'abord l'IA qu'on dit « faible » parce que bien qu'elle soit d'une puissance extraordinaire, elle ne pense pas à proprement parler. Pour être plus précis, on dira qu'elle ne dispose pas de la conscience de soi, de sorte qu'elle se contente pour l'essentiel de résoudre des problèmes à l'aide d'algorithmes qui permettent de traiter d'énormes masses de données (le fameux big data). Ce qui est malgré tout fascinant, ce sont les résultats auxquels elle parvient et les retombées potentielles qui sont les siennes. Qui aurait parié un centime au début du siècle dernier sur le fait qu'une machine battrait aux échecs le meilleur joueur du monde ? Depuis la victoire de l'ordinateur « Deep Blue » contre le champion du monde d'échec, Garry Kasparov, en 1997, nous savons non seulement que c'est possible, mais désormais, une simple application sur un smartphone est capable d'en faire autant. Depuis, grâce à ce qu'on appelle le « deep learning », l'apprentissage profond qui, pour dire les choses simplement, permet à la machine d'apprendre toute seule, en permanence, sans jamais s'arrêter de progresser, l'IA a réussi en 2016 à battre le champion du monde de jeu de GO, ce qui est infiniment plus compliqué que les échecs parce qu'il s'agit d'un jeu dans lequel on ne

peut pas tout calculer. Une telle performance ne pouvait pas rester sans conséquences dans des secteurs de l'existence humaine autrement plus importants, à commencer par la médecine, le droit, l'économie, l'architecture, le bâtiment, la finance et, d'une manière plus générale, le monde du travail, tous les métiers, ou peu s'en faut, étant impactés, parfois de manière mortelle, par les progrès faramineux de l'IA faible. Par exemple, en radiologie, la machine s'est montrée bien supérieure cette année même dans une compétition qui l'opposait à 52 dermatologues provenant de 17 pays différents à qui l'on demandait de poser des diagnostics à partir de photos de grains de beauté. L'IA faible séquence le génome, lit des radios mieux qu'un radiologue, elle commence à faire des merveilles dans le domaine de la traduction automatique, elle gère des pans entiers de l'économie nouvelle avec des applications comme Uber ou Airbnb qui fondent l'économie dite « collaborative », une économie qui se caractérise par le fait que l'intelligence artificielle permet dans nombre de secteurs à des non-professionnels de concurrencer les professionnels de la profession. Elle aura des retombées majeures dans le domaine de la défense, l'organisation du trafic routier ou aérien, la surveillance à domicile des personnes très dépendantes, la lutte contre la criminalité et le terrorisme, l'organisation des secours humanitaires, et dans 1 000 autres secteurs encore. À vrai dire, presque aucun domaine ne sera épargné dans le monde du travail. Ce qui est sidérant, c'est que la vieille Europe commence tout juste à s'en rendre compte alors que la recherche se développe de manière exponentielle aux États-Unis et en Chine depuis déjà plusieurs décennies.

Le deuxième visage de l'IA est celui de la « super IA », qui reste encore une IA faible, mais qui serait contextualisante. Le propre de l'IA faible encore aujourd'hui, c'est en effet qu'elle est assez peu capable de contextualiser les demandes qu'on lui adresse. Elle peut battre le champion du monde de jeu de Go ou séquencer le génome d'une tumeur, mais sortie de son « couloir », elle ne comprend rien. Son intelligence et ses performances sont « verticales », mais fort peu « horizontales » de sorte qu'il lui faut parfois des efforts considérables pour trouver une solution là où un enfant de 5 ans réussit mieux ►►



qu'elle. L'idée qui anime les chercheurs en IA aujourd'hui, c'est qu'il faut la rendre plus transversale et plus contextualisante de telle manière qu'elle sorte de son couloir et devienne, selon la définition que Nick Bostrom donne de la « super IA », « supérieure à l'être humain, non pas dans un seul domaine comme les échecs, le jeu de Go ou le séquençage du génome, mais dans tous les domaines ». Pour prendre encore un exemple que me donnait Nick Liden, le patron de Google Europe, il faut que votre smartphone comprenne que quand vous demandez l'adresse d'un restaurant chinois à 12h et que vous faites la même demande à 3h du matin, la réponse doit être différente. Question de contexte en effet : dans un cas vous voulez sans doute aller déjeuner, dans l'autre vous cherchez plutôt une livraison à domicile. La super IA n'est pas encore au point, mais n'en doutez pas, en Chine et dans la Silicon Valley, on y travaille presque jour et nuit...

Le troisième visage de l'IA serait (je mets au conditionnel, car il s'agit encore d'une utopie) celui de l'IA forte, une intelligence dotée, comme la nôtre, de conscience de soi, de libre arbitre et d'émotions, mais incarnée (si l'on peut dire) sur une base de silicone et non plus de carbone. Alors, nous aurions créé une post-humanité dont nous deviendrions, comme le pense Elon Musk, les animaux domestiques. Pour aller à l'essentiel, on pourrait dire que l'IA forte serait l'intelligence d'une machine capable, non pas seulement de mimer de l'extérieur l'intelligence humaine, mais qui serait bel et bien dotée de trois éléments jusqu'à présent exclusivement humains : la conscience de soi, la faculté de prendre des décisions et les émotions (l'amour et la haine, la peur, la souffrance et le plaisir, la jalousie, etc.). Elon Musk, entre autres brillantissime inventeur de Paypal et patron de Tesla, Space X, Hyperloop ou Neuralink, a déclaré un jour que c'était la plus grande menace qui pèse aujourd'hui sur l'humanité, et joignant le geste à la parole, Musk a mis de sa poche 10 millions de dollars dans un fonds dédié à la recherche sur la sécurité des futures avancées de l'intelligence artificielle, montrant ainsi, s'il en était encore besoin, que l'idéal de la régulation est peut-être bien vital pour nous, aujourd'hui. Certains sont convaincus que l'IA va bouleverser nos repères et

nos valeurs. Mais de quels repères et de quelles valeurs parlons-nous ? Des valeurs des athées ou des valeurs religieuses traditionnelles des juifs, des chrétiens, des musulmans, des bouddhistes ? Des valeurs des socialistes ou de celles des libéraux, de celles des aristocrates ou des démocrates, de l'extrême gauche ou de l'extrême droite, des valeurs qui

**« NOUS NOUS SOMMES CONDUITS EN
PERMANENCE COMME DES MONSTRES.
ON A ENFILÉ LES GÉNOCIDES LES
UNS APRÈS LES AUTRES. »**

furent portées par le nazisme, le stalinisme ou les djihadistes islamistes d'aujourd'hui ? La question semble postuler que les humains auraient des valeurs à la fois communes et magnifiques, des repères fondamentaux que l'intelligence artificielle pourrait venir abîmer, mais la réalité du XX^e siècle vous montre qu'il n'en est rien et que nous nous sommes conduits en permanence comme des monstres. On a enfilé les génocides les uns après les autres, depuis celui des Arméniens jusqu'à celui des Tutsis, en passant par ceux des juifs ou des Khmers et ce fut quasiment sans interruption de sorte que j'ai quelque peine à voir ce que l'IA pourrait venir nous faire perdre de si précieux... Bien entendu, il faudra réguler ce nouveau monde de la technique. On ne pourra ni tout autoriser ni tout interdire, mais ce sera difficile, car la technoscience nous échappe sans cesse davantage pour trois raisons de fond : elle progresse vite, très vite, de sorte que les politiques arrivent toujours en retard ; elle est en outre très difficile à comprendre et surtout, elle est mondialisée, de ce fait les législations nationales n'ont plus grand sens. Seule une prise de conscience européenne, voire mondiale, pourrait avoir quelque efficacité. Il est temps que le monde politique et le monde intellectuel sorte enfin de leur ignorance crasse face à ces sujets... ■