



HAL
open science

L'intelligence de l'intelligence artificielle

Boris Barraud

► **To cite this version:**

Boris Barraud. L'intelligence de l'intelligence artificielle. Boris Barraud. L'intelligence artificielle – Dans toutes ses dimensions, L'Harmattan, 2019. hal-02327501v2

HAL Id: hal-02327501

<https://hal.science/hal-02327501v2>

Submitted on 16 Nov 2019 (v2), last revised 10 Jan 2020 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

L'intelligence de l'intelligence artificielle

Boris Barraud, in *L'intelligence artificielle – Dans toutes ses dimensions*,
L'Harmattan, coll. Europe & Asie, 2019

7 novembre 2019. *L'intelligence artificielle découvre
que la Terre tourne autour du Soleil*¹.

En des temps immémoriaux, Zeus confia à Épiméthée, fils de Japet et Clymène, le soin de distribuer aux espèces vivantes les qualités et les vertus en proportions équilibrées (aux uns les griffes et les crocs pour chasser, aux autres les sabots, les écailles ou les piquants pour se défendre). Épiméthée s'acquitta fort bien de sa tâche et tous les animaux se trouvèrent avec des forces propres leur permettant de jouer un rôle sur la grande scène de la vie. La nature pouvait prospérer et se pérenniser. Mais Épiméthée oublia l'homme². Toutes les capacités et tous les attributs étaient épuisés, il ne restait rien pour l'être humain. Ému de voir ce dernier nu et désarmé, Prométhée, le frère d'Épiméthée, déroba alors à Héphaïstos et Athéna le feu créateur de tous les arts et l'offrit au plus misérable des animaux. Et ce larcin lui valut une lourde punition de la part des dieux, car le monde équilibré conçu par Épiméthée venait de basculer dans la tyrannie : une espèce unique, parce qu'elle serait la seule à maîtriser la technique, allait régner sur l'univers et triompher de tous les autres animaux.

Le mythe d'Épiméthée, raconté par Platon dans le *Protagoras*³, est riche d'enseignements. Bernard Stiegler y voit un rappel de l'état de détresse ou, comme disent les biologistes, de prématuration dans lequel se trouve initialement l'homme, mais ce don divin qu'est la technique lui permet de remédier à ses carences et à sa fragilité originelles⁴. En raison de l'intervention

¹ Communiqué publié dans la revue *Nature* le 7 novembre 2019, indiquant que « les réseaux de neurones peuvent être utilisés pour découvrir des concepts physiques sans aucune connaissance préalable ». L'information peut prêter à sourire, le principe de l'héliocentrisme étant connu depuis les travaux de Nicolas Copernic, exposés en 1533. Plus exactement, une IA a découvert seule que la Terre tourne autour du Soleil. Un réseau de neurones, surnommé « SciNet », fonctionnant à partir de données expérimentales, donc de simples observations objectives, a compris qu'il fallait placer le Soleil au centre du système solaire en fonction des mouvements de l'étoile et de la planète Mars vus depuis la Terre. Selon ses auteurs, la portée de l'étude en question, qui doit paraître plus tard dans les *Physical Review Letters*, est que de telles IA pourraient à l'avenir découvrir de nouvelles lois de la physique, notamment en mécanique quantique qui connaît aujourd'hui beaucoup de contradictions, avant les hommes — mais en réalité, ces IA étant des outils fabriqués par les hommes, ces découvertes hypothétiques seraient plus exactement réalisées par des hommes au moyen de leurs outils les plus perfectionnés.

² En grec ancien, « Ἐπιμηθεύς / *Epimêtheús* » signifie « qui réfléchit après coup ».

³ Platon, *Protagoras*, vers 390 av. J.-C., 320d-322a.

⁴ B. Stiegler, *La technique et le temps – La faute d'Épiméthée*, Galilée, 1994.

providentielle de Prométhée, la faiblesse congénitale de l'espèce humaine s'est transformée en supériorité écrasante, en puissance étouffante, en pouvoir destructeur — pour les autres êtres vivants, mais peut-être aussi pour lui-même.

Les outils des hommes prolongent leurs corps et expriment leurs esprits. Comme l'enseigne l'anthropologue André Leroi-Gourhan⁵, chaque instrument extériorise une fonction bien déterminée. Furent concernés tout d'abord les usages matériels les plus simples et durs (avec le bâton ou le marteau). Puis les apports de la technique devinrent de plus en plus sophistiqués et révolutionnaires (avec les engins énergétiques décuplant ou remplaçant de façon beaucoup plus efficace la force brute des muscles : brouette, poulie, machine à vapeur, chemin de fer, automobile etc.). Plus tard, vinrent les instruments ou machines sensoriels affinant et développant les sens et la perception (lunette, microscope, télégraphe optique, photographie, cinéma etc.). Enfin, apparurent les systèmes informationnels, la dernière et la plus actuelle des générations de la technique. Le but est cette fois de démultiplier et même subroger les fonctions et capacités intellectuelles des hommes grâce aux livres, calembres, ordinateurs, disques durs, réseaux électroniques et informatiques etc. Mais déjà la décomposition alphabétique, la logique, la rhétorique ou les arts de la mémoire ont constitué autant d'avancées techniques immatérielles, toujours utilisées aujourd'hui, permettant de seconder et améliorer l'effort millénaire d'analyse et de synthèse de données ou de connaissances, façonnant en profondeur les habitudes intellectuelles de « penser/classer »⁶.

Les organes technologiques et prothèses mécaniques ou électroniques ont accompagné de près l'homínisation (processus par lequel les primates se sont transformés en hommes) et, surtout, l'humanisation (processus par lequel les hommes préhistoriques sont devenus des hommes historiques, les prémodernes des modernes, les sauvages des civilisés, les obscurs des lumières, les individus des citoyens, les hordes barbares des sociétés paisibles et les regroupements anarchiques des organisations étatiques). Mais se pourrait-il qu'un jour l'association progressiste des hommes et de leurs machines atteigne son point culminant, ne permette plus d'avancées de la civilisation, entraîne des régressions, stimule des phénomènes de déshomínisation et de déshumanisation car l'appareillage technologique serait devenu par trop envahissant, intimidant, inhibiteur et sclérosant. Et ce jour pourrait-il appartenir déjà au passé ?

Quel qu'en soit le sens, l'évolution biologique est donc intimement liée à l'évolution technique. La biosphère est fortement, et de plus en plus, dépendante de la technosphère. L'influence de l'homme sur la technique est quasi-totale ; mais l'influence de la technique sur l'homme est également très importante. Les relations entre l'humain et l'outil, l'être et la machine, le vivant et le robot, le naturel et l'artificiel jouent bien le rôle principal dans le grand film de l'humanité, de l'histoire, des sociétés et des cultures. Les uns créent et utilisent les autres qui, en retour, rétroagissent sur eux, les influencent, modèlent leurs pensées et leurs actes. On peut de moins en moins séparer les êtres humains et les sociétés de leur environnement technique et technologique. Des premiers silex taillés par l'homme préhistorique à l'invention de l'écriture, de Gutenberg à la numérisation massive des données, la longue histoire des techniques et des technologies, aussi ancienne que l'histoire elle-même — et qui en est d'ailleurs à l'origine —, se situe au cœur de l'évolution humaine. Les instruments que l'homme invente le transforment, modifient ses activités personnelles et professionnelles, ses moyens de communication, ses manières de consommer, ses modes de transport ou encore ses façons d'observer le monde et de le penser. Et plus le temps passe, plus l'intimité problématique entre l'homme et la technologie devient forte, à tel point qu'il arrive qu'on ne sache plus si l'on se trouve face à l'un ou à l'autre — et alors le « test de Turing » est passé avec succès.

⁵ A. Leroi-Gourhan, *Le geste et la parole*, Albin Michel, 1964.

⁶ G. Percec, *Penser/Classer*, Hachette, 1985.

Tel est notamment le cas lorsqu'on échange avec des intelligences artificielles, avec des algorithmes, avec des robots dont les réponses atteignent un tel degré de pertinence qu'elles semblent avoir été produites par un cerveau naturel.

L'histoire des techniques est rythmée par le franchissement successif de différents seuils : les cliquets d'irréversibilité. L'invention de l'imprimerie par Gutenberg a ainsi signifié la disparition des moines copistes du Moyen Âge, la création des télégraphes optiques et des chemins de fer a rendu obsolètes les chevaux, les attelages et les cochers, l'introduction de l'informatique et des calculatrices électroniques a sonné le glas des bouliers et des opérations gribouillées sur le papier, l'apparition des fichiers numériques et du *streaming* a conduit à remiser au placard les lecteurs CD et les chaînes hi-fi, qui eux-mêmes avaient entraîné la disparition des gramophones d'antan et la marginalisation des disques vinyles. Nul doute que l'intelligence artificielle constitue un nouveau — et très puissant — cliquet d'irréversibilité.

D'ailleurs, les révolutions de l'intelligence artificielle n'en sont peut-être qu'à leurs prémices. Nul ne saurait dire quelles en seront les conséquences dans 20, 50 ou 100 ans. C'est ce qui rend le sujet passionnant, restant encore largement à explorer, à défricher, à préciser. Et c'est en même temps ce qui suscite une part de fantasmes et de craintes. Les futurologues, mais aussi les spécialistes des nouvelles technologies, l'affirment : l'intelligence artificielle va modifier, souvent radicalement, beaucoup de pratiques. Leurs divergences sont en revanche grandes quant au tour que prendront ces pratiques chamboulées. Or les chercheurs seront les premiers concernés : leurs objets d'étude vont largement évoluer ; et leurs moyens de les étudier vont fortement se renouveler. *Dans toutes ses dimensions*, l'intelligence artificielle est un sujet d'actualité et d'avenir pour les scientifiques de toutes les disciplines.

Les dictionnaires de la langue française définissent l'intelligence comme « la fonction mentale d'organisation du réel en pensées chez l'être humain et en actes chez l'être humain et l'animal »⁷. Il s'agit d'un « ensemble de fonctions psychiques et psycho-physiologiques concourant à la connaissance, à la compréhension de la nature des choses et de la signification des faits »⁸. Ainsi entendue, l'intelligence de l'intelligence artificielle est nulle ou très faible — on y reviendra.

Mais l'intelligence est aussi « l'acte d'appréhender quelque-chose par l'intelligence, de le comprendre avec aisance, d'en avoir la connaissance approfondie »⁹. Sous cet angle, l'intelligence de l'intelligence artificielle est encore balbutiante et très incomplète, ce qui invite à y contribuer, à participer à son approfondissement et à sa précision. L'objet de ce livre est ainsi de proposer une contribution à l'intelligence de l'intelligence artificielle en regroupant les textes de certains des meilleurs spécialistes de l'IA. Il existe un accord parmi les experts sur le fait que l'intelligence artificielle s'est déployée de façon pragmatique, à mesure que des recettes « qui marchent » ont été découvertes, mais on manque encore de concepts et de théories pour la comprendre, la cerner et l'expliquer¹⁰. En ce sens, le besoin d'intelligence de l'intelligence artificielle serait bel et bien réel. Jusqu'à la définition de l'intelligence artificielle, à base de critères et sous-critères permettant de la reconnaître et de délimiter son périmètre, semble encore faire défaut.

L'intelligence artificielle paraît lointaine, justement car on saisit mal ce que cette expression *a priori* oxymorique désigne précisément. L'IA touche pourtant chacun dans son quotidien et dans son intimité. Elle présente également un côté magique, parce qu'elle produit des résultats impressionnants sans qu'il soit possible, pour la plupart des observateurs non mathématiciens et non informaticiens, de comprendre comment elle procède, quels sont ses mécanismes. Une définition de

⁷ V° « Intelligence », in *Trésor de la langue française*.

⁸ *Ibid.*

⁹ *Ibid.*

¹⁰ La situation est ainsi inverse de celle de l'informatique qui tente de mettre en œuvre les concepts définis par des théoriciens parfois fort anciens.

base et quelques premières explications sont donc nécessaires, à titre d'introduction générale du présent ouvrage.

Une intelligence artificielle est une technologie capable de produire des résultats similaires à ceux issus du cerveau humain. Il s'agit d'un outil informatique¹¹ qui effectue des actions ou exécute des tâches qui, il y a peu, étaient le propre des êtres vivants, humains ou animaux¹². Cet outil repose notamment sur des algorithmes, c'est-à-dire des suites de formules mathématiques et de traitements statistiques. Il fonctionne avec des « entrées » (les données initiales) et aboutit à des « sorties » (les résultats) en suivant différentes étapes qui requièrent des calculs, des opérations logiques, des comparaisons ou des analogies. L'intelligence artificielle désigne donc les dispositifs technologiques visant à simuler et, *in fine*, remplacer l'intelligence naturelle, cherchant à reproduire les capacités de l'homme et de l'animal à percevoir, discerner, comprendre, apprendre, raisonner, calculer, mémoriser, comparer, choisir etc. Mais ce que l'IA cherche à imiter, ce n'est pas le fonctionnement de l'intelligence biologique mais uniquement les résultats auxquels elle parvient — le rapport est un peu le même que celui de l'avion avec le vol des oiseaux ou celui de la voiture avec la course à pied. L'ambition des chercheurs en la matière n'est pas et n'a jamais été de fabriquer des cerveaux technologiques se rapprochant au plus près des cerveaux naturels, de la conscience et de la pensée humaines ; elle est uniquement de concevoir des outils informatiques capables de produire les résultats les plus pertinents, essentiellement grâce à du traitement statistique de données.

« Intelligence » dérive du latin « *intelligentem* » ou « *intelligere* », signifiant comprendre, discerner. « *Legere* » veut dire « choisir » et le préfixe « *inter* » se traduit par « entre ». Est donc intelligent ce qui a la capacité de choisir entre différentes alternatives et, pour ce faire, de relier les choses entre elles afin de mieux les saisir. Les hommes et les animaux possèdent certainement cette faculté. L'aspect « intelligent » de l'intelligence artificielle correspond dès lors au but et à la faculté d'imiter le comportement humain ou animal, tandis que l'aspect « artificiel » se rapporte à l'usage d'ordinateurs et autres procédés électroniques et technologiques. L'intelligence artificielle emprunte ainsi à l'informatique, à la logique, aux mathématiques, à l'électronique, aux sciences cognitives, aux sciences de la communication et même à l'éthologie et à la neurobiologie — avec les techniques d'apprentissage profond inspirées des réseaux neuronaux interconnectés, que l'on présentera plus tard dans cette introduction.

Cependant, si une IA peut être capable de restructurer son système de connaissances ou de perception des données à l'aune de nouvelles informations reçues, aucun ordinateur, aucun algorithme ni aucun réseau de neurones ne pense ni n'est intelligent à proprement parler — sa capacité d'apprentissage et de décision est programmée ou, du moins, automatisée, donnant seulement une illusion d'intelligence, au mieux une intelligence faible car très mécanique. Une IA ne peut en effet décider en toute autonomie de ses buts ni concevoir de nouvelles représentations du monde qui la guideraient dans ses raisonnements, qui d'ailleurs n'existent pas puisqu'il ne s'agit que de réflexes et d'automatismes. C'est pourquoi l'appellation « intelligence artificielle » est largement excessive. *De facto*, il ne s'agit guère d'intelligence mais seulement de calculs (certes savants), de statistiques et d'opérations informatiques.

Parler d'« intelligence artificielle » est donc un abus de langage. « Intelligence artificielle » vient de l'anglais « *artificial intelligence* ». Cette expression désigne dans la langue de Shakespeare la simple gestion de données, le traitement de l'information, et non l'intelligence artificielle telle qu'on la comprend en français — qui se traduirait plutôt par « *artificial cleverness* », désignant le

¹¹ L'informatique désigne le traitement rationnel, notamment par des machines automatiques, de l'information.

¹² Marvin Minsky, l'un des scientifiques qui ont créé l'expression « intelligence artificielle », définissait l'IA comme « la construction de programmes informatiques qui s'adonnent à des tâches pour l'instant accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau tels que l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique ».

sens commun, l'empathie, l'arbitraire éclairé, la faculté de comprendre sa propre intelligence et d'en avoir conscience. En somme, il n'est pas déceptif, en anglais, d'évoquer une « *artificial intelligence* » afin de désigner ces nouvelles technologies. Mais il est maladroit, en français, de parler d' « intelligence artificielle ».

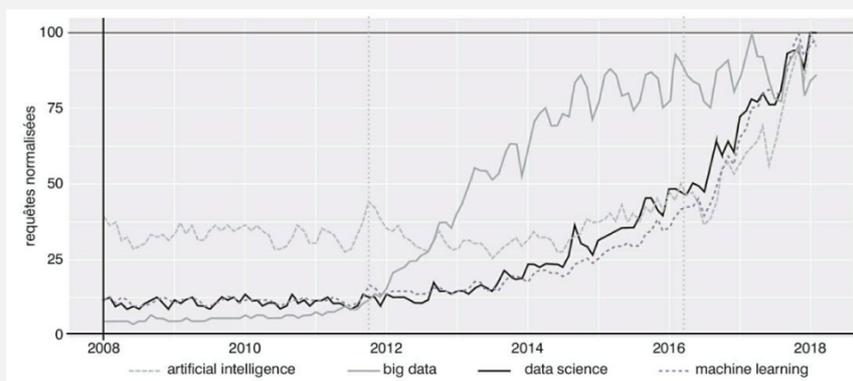
À l'heure actuelle, les robots ne raisonnent pas, ne croient en rien, ne prennent guère de décisions arbitraires, n'ont pas d'intentions ni d'attentions, sont incapables de se donner des buts ou des objectifs nouveaux, n'organisent pas le réel en pensées, sont incapables d'élaborer un concept ou d'avoir une idée, n'ont pas de culture et encore moins de culture générale, n'ont pas conscience d'eux-mêmes et ne possèdent aucun instinct de survie — autant de marques de l'intelligence. Les hommes se caractérisent par l'esprit, par la liberté et par la capacité à s'ouvrir à l'inconnu. Les robots en sont tout le contraire. Pourtant, parler d' « intelligence artificielle » revient dans une large mesure à parler d' « humanité artificielle ». « Intelligence artificielle » est un anthropomorphisme. On humanise, on vitalise ce qui ne sont toujours que des objets animés, des automates, des supercalculateurs¹³. Peut-être vaudrait-il mieux évoquer l' « apprentissage automatique », des instruments « statistiques » ou « prédictifs » ou bien des « agents autonomes » ou « adaptatifs ». L'IA examine aujourd'hui un scanner mieux qu'un radiologue, établit un diagnostic en cancérologie mieux qu'un cancérologue, conduit une automobile mieux qu'un chauffeur, mais elle ne sait pas ce qu'est une radiographie, ce qu'est un cancer, ce qu'est une automobile ni ce en quoi consiste la conduite.

L'expression « intelligence artificielle » (ou « IA ») s'étant partout imposée, tant dans les débats grand public que parmi les discussions de spécialistes — et même comme nom d'un nouveau domaine de recherches scientifiques —, il est cependant difficile de désigner l'objet de ce livre par d'autres mots, sous peine de ne pas se faire comprendre. Par ailleurs, comme il s'agit d'une expression-valise, désignant des technologies disparates et variées, il serait logique de se référer plutôt aux « intelligences artificielles » (au pluriel). Mais l'usage privilégie le singulier dès lors qu'il s'agit de la notion générique d'intelligence artificielle.

La mode de l' « intelligence artificielle » et de son sigle « IA » est récente, comme en témoignent les statistiques de requêtes sur le plus célèbre moteur de recherche du web. Cette mode n'en est pas moins forte et, si l'on veut parler de l'IA, on ne peut que parler d' « IA ».

¹³ Yann Le Cun estime toutefois que l'IA pourrait bientôt exprimer des sentiments et des émotions. Mais il précise bien que ce ne serait que parce qu'il deviendrait possible qu'un cerveau artificiel calcule la joie, la peur ou la tristesse. Cependant, de telles joie, peur ou tristesse calculées et non ressenties ou éprouvées en seraient-elles vraiment ?

RECHERCHES INTERNET SUR LES SUJETS RELIÉS À L'INTELLIGENCE ARTIFICIELLE



— Source: d'après Google Trends. Les données sont normalisées. Pour chaque requête, la normalisation est faite en pourcentage du nombre maximal de requêtes observées sur la période.

L. Galiana, « Les conséquences économiques de l'intelligence artificielle », *Idées économiques et sociales* 2018, n° 192, p. 27

C'est en 1956 que les mathématiciens américains John McCarthy, Marvin Minsky, Claude Shannon et Nathaniel Rochester organisèrent, au Dartmouth College, dans l'État du New Hampshire, un séminaire sur les « machines pensantes » au cours duquel ils proposèrent pour la première fois de parler d'« *artificial intelligence* »¹⁴. Un nouveau domaine d'enseignement et de recherche scientifiques était né, dont l'objectif, longtemps et aujourd'hui encore perçu comme farfelu ou irréaliste, est de réaliser des machines capables d'apprendre — notamment de leurs erreurs —, de comprendre leur environnement et le monde en général, de s'adapter aux changements, d'anticiper et de prévoir le futur, d'agir afin d'améliorer la condition humaine. Plus cet objectif initial sera atteint et moins il sera décevant de se référer à l'« intelligence » artificielle. On pourra alors comparer à raison cette dernière et l'intelligence biologique.

Pour l'heure, néanmoins, ce qu'on appelle « intelligence artificielle » demeure un simple produit de l'intelligence humaine, sa création et son reflet, globalement immensément inférieur à elle. Pour devenir une réalité, l'intelligence artificielle devra gagner en autonomie, en libre-arbitre, se détachant des hommes qui l'ont faite naître. Il lui faudra devenir une intelligence artificielle générale, capable de comprendre le fonctionnement du monde dans sa globalité par simple observation, comme le fait un enfant durant les premières années de sa vie. Or il lui sera difficile d'y parvenir¹⁵. On imagine difficilement l'intelligence artificielle rejoindre et dépasser un jour l'intelligence naturelle autrement que ponctuellement, concernant des tâches très précises susceptibles de reposer sur des traitements statistiques. Beaucoup de décisions supposent des choix arbitraires plutôt que la simple poursuite de tendances statistiques.

¹⁴ On attribue également la paternité, si ce n'est de l'expression, de la pensée de l'intelligence artificielle au mathématicien britannique Alan Turing, qui s'y est intéressé dès 1950 à l'occasion de ses travaux sur le « jeu de l'imitation » (ou « test de Turing ») : ce test consiste à mettre un humain en confrontation verbale à l'aveugle avec un ordinateur et un autre humain ; si cette personne n'est pas capable de dire lequel de ses interlocuteurs est un ordinateur, alors l'ordinateur a passé avec succès le test (A. Turing, « *Computing Machinery and Intelligence* », *Mind* oct. 1950).

¹⁵ Bien qu'on commence à appliquer le *machine learning* à lui-même, donc à entraîner des ordinateurs à concevoir des réseaux de neurones, si bien que des intelligences artificielles pourraient un jour créer d'autres intelligences artificielles.

Lorsque les chercheurs du Dartmouth College créèrent l'expression « *artificial intelligence* », il s'agissait d'ailleurs surtout d'une démarche marketing, presque un mensonge volontaire : recourir à une expression sensationnelle afin d'obtenir des crédits de recherche de la part des pouvoirs publics américains¹⁶. Le concept d'intelligence s'attache à un fond ontologiquement biologique et l'intelligence « artificielle » ne peut pas être intelligente, par définition. Mais cela n'empêche pas l'IA de rendre bien des services tout en posant de nombreux défis à l'intelligence humaine, en premier lieu celui de son intelligence, c'est-à-dire celui de sa compréhension, de son explication et de son analyse critique.

L'intelligence de l'intelligence artificielle suppose de combiner quatre efforts cognitifs afin de mieux appréhender cette méta- et méga-révolution (source de révolutions nombreuses et variées) :

- accepter que l'IA n'est qu'un outil développé par et pour les hommes — point de vue humain (I) ;
- s'intéresser aux rouages et au fonctionnement de la technologie — point de vue technique (II) ;
- se concentrer sur les conséquences civilisationnelles de l'IA — point de vue humain, social, politique et économique (III) ;
- donner la parole aux chercheurs et l'écouter — point de vue scientifique (IV).

I. Une problématique humaine : l'intelligence de la machine au service de l'intelligence de l'homme

Des robots mus par des IA en venant à défier et affronter les hommes est de l'ordre du pur fantôme. Présenter le rapport entre les uns et les autres sous les traits d'un duel dont l'un des protagonistes devrait triompher de l'autre ne paraît guère réaliste. Les humains produisent des robots qui ne sont et ne seront que leurs outils ; et qui, comme tous outils, ne sont ni bons ni mauvais en soi car soumis aux destinations qu'on leur attribue, aux usages qu'on en fait¹⁷. Il n'y a et n'y aura pas plus de compétition entre l'homme et l'IA qu'entre le pianiste et le piano ou entre le maçon et la truelle. Et puis il serait difficile de dire qu'une espèce aussi proliférante que l'espèce humaine pourrait s'effacer et perdre la mainmise sur le monde, surtout au profit de ses propres technologies.

En 2014, l'astrophysicien Stephen Hawking affirmait pourtant, dans une interview à la BBC, que « le développement d'une intelligence artificielle complète pourrait provoquer la fin de l'humanité »¹⁸. Il s'inquiétait d'éventuels périls liés à l'avènement de robots qui surpasseraient des

¹⁶ Un travestissement de la réalité qui rappelle celui d'Éric Le Rouge lorsqu'il emmena ses hommes au Groenland en leur faisant croire que c'était un « pays vert » (traduction de « *groenland* »), alors que c'était un pays blanc, recouvert de neige et de glace. Mais le nom « Groenland » est resté.

¹⁷ C'est là l'occasion de rappeler les lois énoncées par l'écrivain de science fiction Isaac Asimov (*Cercle vicieux*, Runaround, 1942) :

« 1° Un robot ne peut porter atteinte à un être humain, ni, en restant passif, permettre qu'un être humain soit exposé au danger ;

2° Un robot doit obéir aux ordres qui lui sont donnés par un être humain, sauf si de tels ordres entrent en conflit avec la première loi ;

3° Un robot doit protéger son existence tant que cette protection n'entre pas en conflit avec la première ou la deuxième loi ».

¹⁸ L'homme ne représente que 0,01 % des êtres vivants mais est à l'origine de 99 % des destructions dans la biosphère. Pourrait-il soudain perdre ce monopole de la destruction et se retrouver lui-même détruit par des êtres supérieurs ? Ce scénario, aujourd'hui, relève entièrement de la science fiction.

êtres humains limités par la lenteur de l'évolution biologique. Et beaucoup, comme Bill Gates ou Elon Musk — qui a créé une fondation pour mesurer et prévenir la dangerosité des IA, laquelle serait le plus grand danger ayant jamais pesé sur l'humanité —, entretiennent une psychose effrayante autour de la prise de pouvoir de machines toutes-puissantes créées par l'homme puis lui échappant¹⁹. C'est un lieu commun de faire le procès des nouvelles technologies, presque un « marronnier », revenant périodiquement sur le devant de la scène depuis des siècles. Il y a toujours eu des prophètes, apocalyptiques ou intégrés²⁰, pour annoncer les ruptures radicales et lourdes de conséquences que la rationalisation scientifique des machines provoquerait. Mais ces déclamations tapageuses sont souvent biaisées ou naïvement critiques. Il est difficile d'imaginer qu'un jour la machine puisse en venir à penser comme l'homme, contre l'homme ou mieux et plus vite que l'homme, si bien que l'être humain, de maître et possesseur du monde sous toutes ses formes, se retrouverait en position d'esclave ou de jouet de robots qui ne seraient plus les fruits de son art mais auraient la capacité de se reproduire et de se perfectionner seuls. Tout cela appartient à la « science » fiction.

Cette « science » fiction diffuse différentes légendes parmi le grand public, notamment celle selon laquelle l'IA devrait se comporter à l'image des hommes et présenter les mêmes qualités et les mêmes défauts démultipliés, les mêmes pulsions et le même instinct renforcés — on fabrique des robots humanoïdes, mais leur « intelligence » et leurs « cerveaux » ne partagent quasiment rien avec l'intelligence et les cerveaux humains. Il y a également le mythe selon lequel il existerait un secret scientifique et technologique de l'intelligence artificielle qui resterait à découvrir, si bien que celui qui le découvrirait pourrait dominer le monde — alors que l'intelligence artificielle progresse par petits (et grands) pas, par étapes, en tâtonnant, incluant en particulier des découvertes dont les tenants et les aboutissants (non seulement les aboutissants) sont rendus publics, permettant à chacun de les réutiliser librement et de les améliorer. Le problème qui se pose plus sérieusement, et déjà actuellement, n'est pas celui de machines asservissant les humains mais celui d'humains asservissant d'autres humains grâce aux machines — un problème politique, juridique, social et économique. Là où il faut et faudra toujours de l'éthique, ce n'est pas dans l'intelligence artificielle mais dans l'intelligence humaine²¹.

Les projets d'intelligences artificielles doivent donc s'orienter vers des programmes réfléchis et contrôlés, non abandonnés à des intérêts économiques, mercantiles et vénaux, livrés à la seule quête de nouveaux marchés et de profits croissants, au besoin en manipulant, en enfermant, en limitant au maximum le sens critique et les valeurs morales et en réduisant le monde à deux ensembles : celui des produits et celui des consommateurs. Il importe d'écouter avec du recul et une oreille critique les discours hypocrites ou naïfs de certaines multinationales de l'IA affirmant que leur but ne serait pas de générer des profits et des monopoles mais de permettre à l'humanité de progresser et d'entrer dans un nouvel âge.

Pourtant, la finalité unique ou ultime des projets d'intelligences artificielles devrait effectivement être de faciliter la vie des hommes et le bon fonctionnement des sociétés, de les protéger des dangers qui les guettent — la maladie, la corruption, la manipulation, l'aveuglement

¹⁹ Des prolongements universitaires de ces questionnements existent. À l'université de Cambridge, le Centre pour l'étude des risques existentiels (*Centre for the Study of Existential Risk*) a été créé afin d'étudier les risques et menaces que les technologies feraient peser sur le futur. Quant à l'Institut pour l'avenir de l'humanité (*Future of Humanity Institute*) de l'université d'Oxford, il se consacre aux questions relatives à l'avenir de l'humanité en donnant une grande place aux nouvelles technologies.

²⁰ U. Eco, *La Guerre du faux*, Grasset, 1985.

²¹ En témoigne l'invention, en 2017, par des chercheurs de l'université américaine de Stanford, d'un algorithme capable de découvrir l'orientation sexuelle des personnes en fonction des traits de leurs visages, alors que tant de programmes de recherche en intelligence artificielle plus utiles pourraient être menés. En témoigne, également, l'intelligence artificielle, créée par Microsoft en 2016, qui devait évoluer sur Twitter aux côtés des internautes humains. En quelques heures, ces derniers l'ont rendue complètement misogyne, homophobe, raciste et même nazie, si bien qu'elle a dû être retirée en catastrophe.

etc. Certains de ces projets permettront de sauver de nombreuses vies, notamment dans le domaine des transports ou dans celui de la santé — avec une circulation automobile bien moins accidentogène et des diagnostics médicaux beaucoup plus rapides, précis et fiables²². Neutre en soi, la technologie peut être mise à profit par des hommes qui le souhaitent pour contaminer, corrompre ou manipuler — par exemple avec les *deep fakes*, qui seraient la première de toutes les menaces liées à l'IA²³. Son utilisation en tant qu'outil de propagande ou de désinformation peut être désastreuse. Comme l'internet et comme de nombreuses révolutions technologiques, l'IA est donc un moyen pour les personnes bien intentionnées comme pour les personnes mal intentionnés d'arriver plus facilement à leurs fins.

Dans une perspective dite « transhumaniste », l'intelligence artificielle peut soutenir l'humain dans sa quête de puissance et de pouvoir, donc d'actions rationnelles, de décisions éclairées et pertinentes, de maîtrise et de domination. Elle peut augmenter ses capacités cognitives. Mais elle devrait demeurer son instrument et ne pas le concurrencer²⁴. D'ailleurs, la volonté de domination des uns sur les autres est davantage liée à la testostérone qu'à l'intelligence ; et l'IA ne produit pas de testostérone. L'instinct et les pulsions de domination sont aussi liés au fait que les hommes sont des animaux sociaux qui ont besoin d'interagir avec leurs semblables, de les influencer, de s'imposer à eux. Cela n'existe pas chez les animaux non-sociaux tels que les orangs-outans. Or l'IA n'a pas d'instinct grégaire, elle ressemble, sur ce plan, plus à l'orang-outan qu'à l'homme.

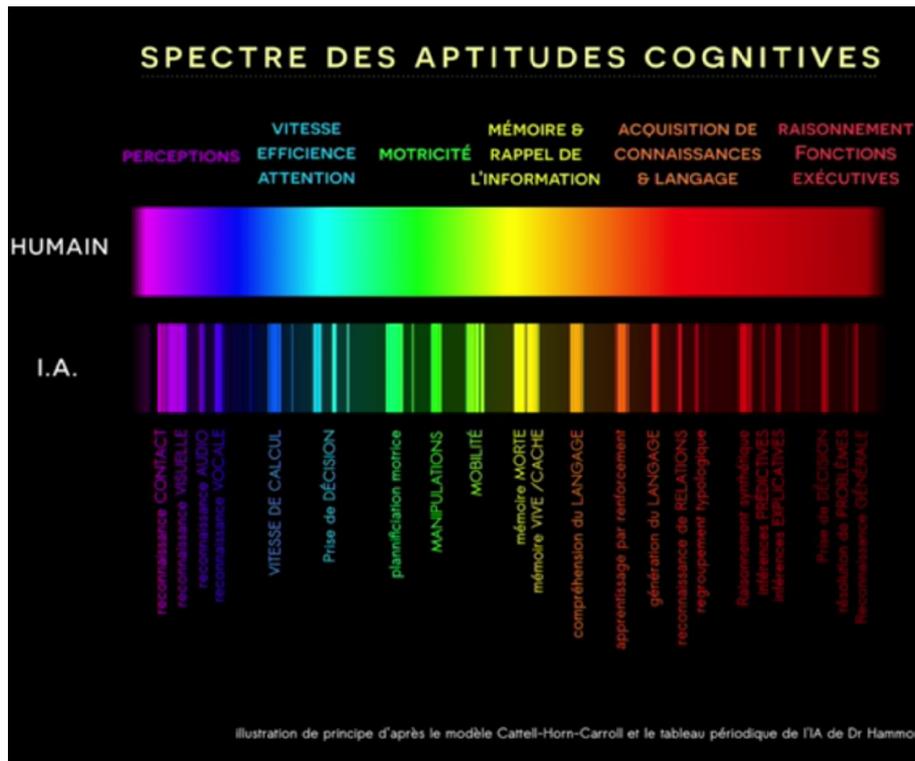
Idéalement, si les humains créent des formes d'intelligence artificielle, ce ne devrait être qu'afin de progresser, de rendre plus aisés nombre de leurs actions individuelles et collectives. L'invention de l'intelligence artificielle se positionne ainsi dans la continuité des inventions du langage, de l'écriture, de l'imprimerie, de la roue, de la machine à vapeur, du moteur à explosion, du système électrique etc. L'homme, conscient des limites de son intelligence, pourtant sans conteste la plus développée de tout le règne animal, se dote de prothèses telles que des bibliothèques, des machines à calculer, des ordinateurs ou des algorithmes. Il parvient ainsi à conserver en mémoire et à trier de grandes masses d'informations ou à réaliser des calculs hyper-complexes en quelques instants.

Cependant, si l'IA est capable de réaliser des prouesses qu'aucun humain ne saurait accomplir, elle ne sait pas faire beaucoup de choses simples pour un homme car relevant du bon sens ou du sens commun, de l'intuition, de l'émotion, de l'empathie, de la créativité etc. L'illustration suivante montre combien les aptitudes cognitives de l'IA sont partielles, bien que parfois très développées, quand les capacités cognitives de l'humain sont totales, transversales, mais limitées par nature dans leurs performances :

²² Dans le diagnostic de certains cancers, l'IA réalise 99 % de bons diagnostics, contre 60 % à 80 % parmi les médecins humains. Ceux-ci sont notamment dépassés pour des raisons physiologiques : par exemple, une rétine humaine, d'un radiologue, distingue 15 niveaux de gris, contre un milliard de niveaux de gris différents perçus par l'IA.

²³ Les *deep fakes* sont de faux documents, notamment vidéos avec des imitations toujours plus réalistes du comportement et de la voix, permettant de travestir la réalité, de faire dire ou faire faire des choses à des individus qui ne les ont jamais dites ni faites, de mettre le visage de quelqu'un sur le corps de quelqu'un d'autre, ou placer une personne dans un contexte dans lequel elle ne s'est jamais trouvée. Ces *deep fakes* sont évidemment très dangereux, car quand on dit que quelqu'un a dit quelque-chose et qu'en plus on en montre la preuve, il est difficile de ne pas y croire. Parmi les scientifiques, ceux qui œuvrent à la création des *deep fakes* sont aussi ceux qui recherchent les moyens de les détecter — avec des algorithmes reconnaissant les vidéos d'origine sur le web, évaluant la cohérence du flux sanguin entre le corps et la tête ou mesurant le caractère naturel du clignement d'œil.

²⁴ Les visions d'une superintelligence informatique vouée à prendre le pouvoir reposent principalement sur la loi de Moore — décrite au milieu des années 1960 par le cofondateur d'Intel Gordon E. Moore, établissant que la puissance des processeurs suit une évolution exponentielle à travers le temps, doublant tous les dix-huit mois. Mais les progrès technologiques vont obligatoirement ralentir et, pour certains, s'arrêter. Si tout n'est pas possible éthiquement et moralement, il faut ajouter que tout n'est pas possible physiquement, scientifiquement, technologiquement.



On considère classiquement, depuis les travaux de psychologie cognitive d'Howard Gardner, qu'il existerait huit formes d'intelligence²⁵. Certaines (visuelle-spatiale, logique-mathématique ou verbale-linguistique) semblent plus déshumanisables que d'autres (interpersonnelle, intrapersonnelle ou naturaliste-écologiste).

LES INTELLIGENCES MULTIPLES

d'après la théorie de Howard Gardner

<p>Intelligence verbo-linguistique</p> <p>ABC LES MOTS</p> <p>Capacité à penser, saisir et exprimer des idées. Exemples : écrivains, poètes, traducteurs, interprètes, journalistes, avocats, ...</p> 	<p>Intelligence musico-rythmique</p> <p>LA MUSIQUE</p> <p>Capacité à reconnaître, interpréter, et créer des musiques, rythmes... Exemples : compositeurs, chanteurs, musiciens, professeurs de musique, accordéons, ...</p> 
<p>Intelligence logico-mathématique</p> <p>123 LA LOGIQUE</p> <p>Capacité à raisonner, calculer, compter, résoudre des problèmes. Exemples : scientifiques, ingénieurs, savants, mathématiciens, comptables ...</p> 	<p>Intelligence interpersonnelle</p> <p>LES AUTRES</p> <p>Capacité à agir et réagir avec les autres, aider, collaborer, partager. Exemples : commerçants, politiciens, enseignants, managers, chef d'équipe, ...</p> 
<p>Intelligence visuelle / spatiale</p> <p>LES IMAGES</p> <p>Capacité à percevoir et se représenter le monde. Exemples : peintres, photographes, architectes, pilotes, dentistes, ...</p> 	<p>Intelligence intrapersonnelle</p> <p>SOI</p> <p>Capacité à se connaître, à s'introspecter, à être intuitif. Exemples : chercheurs, entrepreneurs, romanciers ...</p> 
<p>Intelligence corporelle-kinesthésique</p> <p>LE CORPS</p> <p>Capacité à utiliser son corps, à s'exprimer physiquement. Exemples : danseurs, athlètes, artisans, chirurgiens, pompiers ...</p> 	<p>Intelligence naturaliste-écologiste</p> <p>LA NATURE</p> <p>Capacité à observer, reconnaître, et classifier la nature. Exemples : explorateurs, botanistes, archéologues, zoologistes, fermiers, ...</p> 

Illustrations : Céline Cidre / www.pour

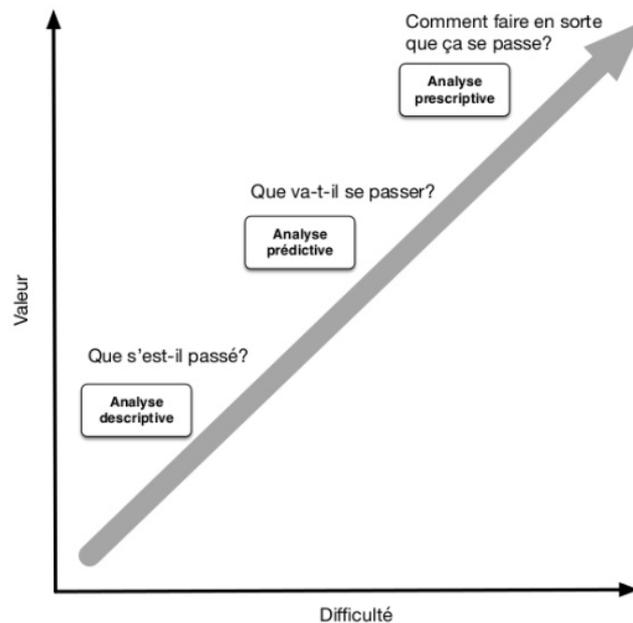
C'est la symbiose de deux formes d'intelligence — ou, dans un cercle vertueux, augmentation de l'intelligence par ses réalisations les plus perfectionnées — qui permet le progrès

²⁵ H. Gardner, *Les formes de l'intelligence*, Odile Jacob, 1997.

des savoirs et des connaissances, ainsi que des idées. Et les chercheurs comptent au nombre de ceux qui en retirent les plus grands bénéfices.

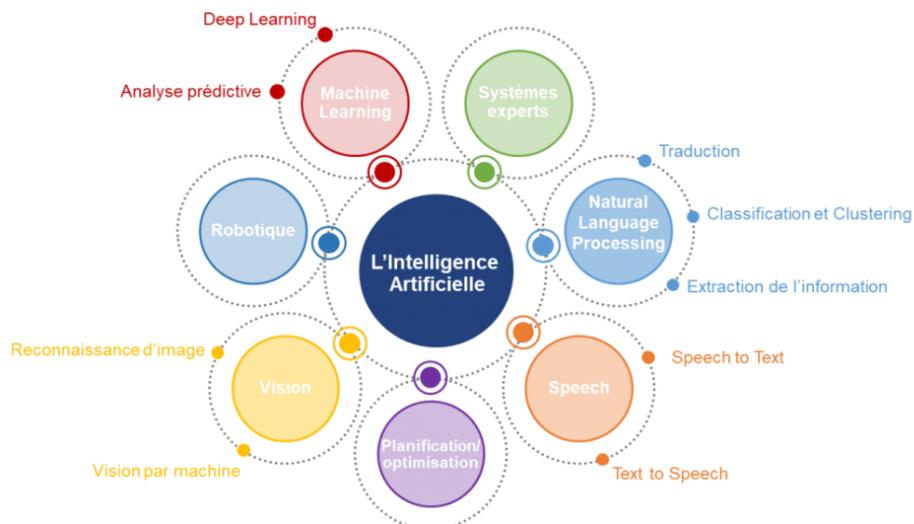
L'intelligence artificielle a vocation à servir l'intelligence humaine en éclairant et donc facilitant ses prises de décisions. Pour ce faire, elle doit lui fournir des informations à haute valeur ajoutée relevant de l'analyse descriptive consistant à trier et croiser les informations. Mais son caractère révolutionnaire tient surtout aux analyses prédictives et même prescriptives dont elle est capable, lesquelles, qualitativement, sont incomparables aux simples analyses descriptives :

Ajouter de la valeur



CETIC, juin 2017

Par ailleurs, les principaux domaines fonctionnels de l'IA témoignent de son immense potentiel, des innombrables réalisations auxquelles elle est susceptible de donner lieu. Il s'agit d'approfondir des domaines fonctionnels déjà connus, à l'instar des systèmes experts, de la planification/optimisation ou de la robotique, mais aussi de développer de nouveaux domaines tels que le *machine learning*, le *natural language processing*, la vision (capacité pour une machine d'appréhender son environnement) ou encore le *speech* (texte vers parole ou parole vers texte).



Panorama des domaines de l'IA

Artik Consulting, juin 2018

Et ces réalisations ont vocation à concerner de très nombreux secteurs :

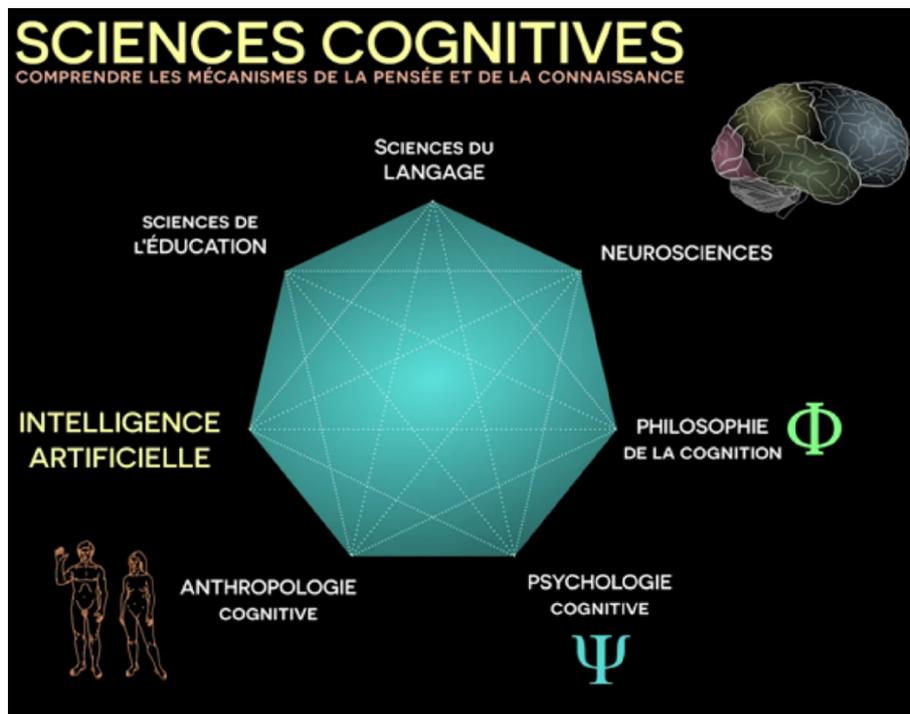
Utilisation des algorithmes et de l'intelligence artificielle

	EDUCATION	JUSTICE	SANTE	SECURITE	TRAVAIL	CULTURE	AUTRES
Générer du savoir	Mieux cerner les aptitudes d'apprentissage des élèves	Mettre en évidence des manières différentes de rendre la justice selon les régions	Tirer profit de la quantité immense de publications scientifiques	Repérer des liens insoupçonnés pour la résolution d'enquêtes par les services de gendarmerie	Comprendre les phénomènes sociaux en entreprise	Créer des oeuvres culturelles (peinture, musique...)	Affiner le profil de risque d'un client d'un assureur
Croiser des infos, «matching»	Répartir les candidats dans les formations d'enseignement (ex APB)		Répartir des patients pour participation à un essai clinique		Faire correspondre une liste de candidatures avec une offre d'emploi		Mettre en relation des profils « compatibles » sur des applications de rencontres
Prédire	Prévoir des décrochages scolaires	Prédire la chance de succès d'un procès et le montant potentiel de dommages intérêts	Prédire des épidémies Repérer des prédispositions à certaines pathologies afin d'en éviter le développement	Détecter les profils à risque dans la lutte antiterroriste Prédire l'occurrence future de crimes et délits	Détecter les collaborateurs qui risquent de démissionner dans les prochains mois	Créer des oeuvres ayant un maximum de chance de plaire aux spectateurs (Netflix)	
Conseiller	Recommander des voies d'orientation personnalisées aux élèves	Recommander des solutions de médiation en fonction du profil des personnes et de cas similaires			Proposer des orientations de carrière adaptées aux profils	Recommander des livres (Amazon), des séries télévisées (Netflix), etc.	Individualiser des messages politiques sur les réseaux sociaux
Aider à la décision		Suggérer au juge la solution jurisprudentielle la plus adéquate pour un cas donné	Suggérer au médecin des solutions thérapeutiques adaptées	Proposer aux forces de police les zones prioritaires dans lesquelles patrouiller			Aider à trouver le chemin le plus court (GPS)

© F. Schneider - Source : CIVIL - Ethique et numérique

II. Une problématique technique : au cœur des cerveaux artificiels (apprentissage automatique et apprentissage profond)

« Intelligence artificielle » peut désigner un objet polymorphe, un champ de recherches relativement vaste ou une science particulière. Dans ce dernier cas, il s'agit de l'une des sciences cognitives qui s'intéressent à la pensée et à l'intelligence :



Audrey Vermeulen, TEDx Canebière, 2018

Dans ce livre multidisciplinaire, on utilisera surtout « intelligence artificielle » afin de désigner un objet et le champ de recherches multifacette auquel il donne lieu. On pourrait d'ailleurs évoquer des « sciences de l'intelligence artificielle » au sujet des disciplines qui étudient cet objet de différents points de vue — sur le modèle des sciences de l'information et de la communication qui comprennent des dimensions multiples, du droit à la sociologie en passant par l'économie. C'est à ces diverses sciences de l'IA que le présent ouvrage à la fois s'adresse et donne la parole. Il considère que cette technologie serait le cœur d'un système dont on pourrait difficilement ne comprendre qu'un des rouages abstraitement de tous les autres. Il mise sur le fait qu'on pourrait difficilement comprendre l'IA seulement en termes techniques, sociologiques, économiques, juridiques, éthiques, philosophiques, anthropologiques ou autres.

En premier lieu, la dimension technique semble la plus incontournable. Si l'IA doit de préférence être conçue tel un objet, il serait bien sûr malvenu de chercher à disserte sur ou autour de lui sans en maîtriser les origines, implications et potentialités technologiques, c'est-à-dire sans bien saisir « comment ça marche ». Le physicien Richard Feynman affirmait en ce sens qu'« on ne peut comprendre véritablement un instrument tant qu'on ne l'a pas construit soi-même ». Tout chercheur, y compris le philosophe, doit aller sur le terrain pour bien comprendre ce autour de quoi il disserte — comme le faisait, par exemple, Adam Smith durant la révolution industrielle. Sans être mathématicien ni informaticien ou *data scientist*, il importe, pour pouvoir en parler, de ne pas se contenter de définitions superficielles ou basiques et de s'intéresser dans le détail au fonctionnement

de la technologie — et pourquoi pas de s'initier au codage informatique. Le chercheur en IA peut difficilement ne pas être technophile, pas dans le sens de technoaddict mais dans celui de passionné par la dimension technique de la technologie, celle-ci expliquant beaucoup de choses sous les divers autres angles.

Historiquement, les origines de l'intelligence artificielle sont très anciennes. Elles s'inscrivent dans la longue histoire des algorithmes, dont les premiers auraient été inventés par les babyloniens il y a 5 000 ans — l'un de ces algorithmes consistant par exemple en une tablette permettant de calculer des racines carrées précises au millionième près. D'autres se retrouvent dans les antiquités égyptienne et grecque. Au XVI^e siècle, Léonard De Vinci conçut nombre d'automates imitant l'homme ou l'animal. Et, au XVII^e siècle, Blaise Pascal inventa la Pascaline, une intelligence artificielle mécanique permettant de réaliser des opérations arithmétiques. Aujourd'hui, l'IA ne peut plus se passer de l'électronique, c'est-à-dire de l'électricité comme support de traitement, de transmission et de stockage d'informations.

Durant les années 1950, alors que l'informatique commençait à progresser à grands pas, l'intention d'élaborer des « machines à penser » se rapprochant dans leur fonctionnement de l'esprit humain s'est affirmée, avec des figures de proue telles qu'Alan Turing. Déjà en 1914, l'ingénieur espagnol Leonardo Torres Quevedo inventait le premier véritable automate joueur d'échecs. Celui-ci était capable de remporter une partie à tous les coups lors d'une confrontation roi et tour contre roi. Prouesse d'une autre envergure, le supercalculateur d'IBM Deep Blue battait, en 1997, le champion du monde Garry Kasparov — qui avait pourtant prédit que « jamais une machine ne pourrait être supérieure à un humain ».

Mais c'est en 2012 que le secteur de l'intelligence artificielle franchit un pallier considérable, sortant de l'« hiver » dans lequel il s'était quelque peu endormi, faute de progrès significatifs. Cette année là, pour la première fois, un réseau de neurones profond remporta le challenge ImageNet, un concours de reconnaissance d'images. Avec un taux d'erreurs de seulement 16,4 %, il distança très largement le second programme (26,2 %). Cette performance remit sur le devant de la scène une technique d'apprentissage profond (*deep learning*) qui avait commencé à faire ses preuves dans les années 1990, notamment avec les travaux du français Yann Le Cun et des canadiens Yoshua Bengio et Geoffrey Hinton, avant d'être un temps laissée de côté. La puissance de calcul des cartes graphiques couplée à la mise à disposition de grandes bases de données étiquetées (une image d'un chien sur un vélo en train de jouer de la guitare est accompagnée des indications « chien », « vélo » et « guitare ») permirent à ces programmes d'enfin faire leurs preuves à grande échelle.

La communauté scientifique perçut alors l'immense potentiel de cette technologie : en étant entraînée sur une base de données suffisamment consistante et pertinente, elle est capable d'établir elle-même des règles et standards d'analyse lui permettant d'interpréter d'autres données jusqu'ici trop complexes à traiter. Dès 2013, seuls des réseaux de neurones profonds participaient au challenge ImageNet. Et, en 2016, AlphaGo battait le maître coréen Lee Sedol au jeu de go, lequel est ô combien plus complexe à maîtriser que le jeu d'échecs²⁶. Vingt ans après les meilleurs joueurs d'échecs, les champions de go, de poker et de nombreux jeux vidéo s'inclinent à leur tour devant des intelligences artificielles — qui ont pour elles les avantages d'être durantes, de ne jamais se déconcentrer, de ne pas subir le stress ni la pression et d'être toujours d'humeur égale, autant de gains psychologiques de la machine par rapport à l'humain. Quant au taux d'erreurs dans la reconnaissance d'images, il a chuté à quelques pourcents. Et des IA sont désormais capables de créer de nouveaux tableaux de Rembrandt ou de nouveaux morceaux des Beatles incroyablement réalistes — bien que les spécialistes ne puissent pour l'heure les qualifier que de « mauvais Rembrandt » ou de « mauvais Beatles ».

²⁶ Quelques mois plus tard, AlphaGo première version a lui-même été battu par son successeur, entraîné sans exemples, uniquement par exploration, sur le score de 100 parties à 0, ce qui témoigne de la vitesse des progrès de l'IA.

Les concepteurs de réseaux de neurones sont un peu des alchimistes : ils associent de diverses manières leurs formules mathématiques et leurs technologies et, lorsqu'elles produisent des résultats satisfaisants, ils cherchent à perfectionner la recette sans nécessairement comprendre son mécanisme exact. En la matière, l'« art de la débrouille » est important. Un réseau de neurones ne se programme pas mais s'éduque. Il s'entraîne et il apprend. Et cet apprentissage peut se faire soit par exemples soit par exploration, c'est-à-dire soit en étant guidé par l'homme soit de façon autonome et automatique. Pour l'heure, c'est l'apprentissage grâce à des exemples et donc des statistiques qui est essentiellement utilisé.

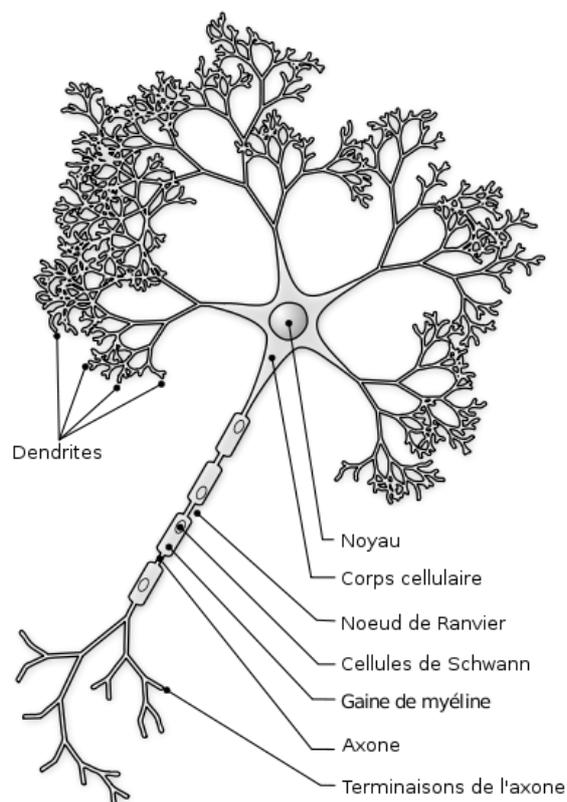
Pour entraîner une intelligence artificielle à reconnaître des chats sur des images contenant aléatoirement des chiens ou des chats, on lui montre un grand nombre d'exemples d'images de chats et un grand nombre d'exemples d'images de chiens. Plus le nombre d'exemples est grand, plus la machine devient capable de reconnaître seule les chats car elle a en mémoire beaucoup de cas de figure possibles avec lesquels elle peut comparer l'image qu'elle traite. C'est cette technique qu'on appelle « apprentissage automatique » (*machine learning*). Lorsqu'on parle d'intelligence artificielle aujourd'hui, c'est très souvent d'apprentissage automatique au moyen de réseaux de neurones et d'exemples qu'il s'agit. Cette intelligence est toutefois bien relative et, tandis que, pour qu'un jeune enfant apprenne à distinguer le chat du chien, il suffit de lui en désigner trois ou quatre de chaque espèce en lui disant « ceci est un chat » et « ceci est un chien », il faut montrer des milliers voire des millions d'exemples à une IA pour qu'elle devienne aussi perspicace.

Le *machine learning* est notamment utilisé en matière de recommandation (rechercher les consommateurs les plus proches d'un individu afin de lui suggérer l'achat de certaines catégories de produits), de *clustering* (définition de groupes typiques, par exemple d'utilisateurs d'un réseau social), de détection automatique de fraudes, anomalies ou erreurs de saisie etc.

Quant au *deep learning* (apprentissage profond), il est le niveau supérieur du *machine learning*. Il désigne des méthodes d'apprentissage dédiées aux phénomènes complexes pour lesquels une structure à plusieurs couches est nécessaire. Il intervient afin de résoudre des problèmes appelant des méthodes plus sophistiquées, comme la reconnaissance automatique de textes, le traitement du langage naturel ou la reconnaissance d'images (analyse des émotions, diagnostics automatiques à partir d'images médicales, compréhension de l'environnement proche pour les voitures autonomes).

Les cerveaux artificiels ne cherchent pas à reproduire le plus exactement possible la constitution et le mécanisme des cerveaux biologiques. Leur objectif est essentiellement d'aboutir aux mêmes résultats, donc d'imiter la nature dans les performances dont elle est capable mais pas dans les mécanismes qui permettent d'aboutir à ces performances. Néanmoins, la technologie des réseaux de neurones artificiels s'inspire du fonctionnement du cerveau. Ce dernier est peut-être l'objet le plus complexe qu'on connaisse dans l'univers et son mécanisme demeure aujourd'hui encore assez mal compris. Il fonctionne grâce à des milliards de neurones. Un neurone est une cellule nerveuse constituant l'unité fonctionnelle du système nerveux. Elle a la particularité de posséder des entrées qui répondent à des stimuli, les dendrites (environ 7 000 par neurone), et une sortie en direction des autres neurones, l'axone. Le neurone reçoit des informations en provenance de milliers d'autres neurones, puis stimule à son tour les neurones voisins en envoyant à travers son axone un signal bioélectrique : l'influx nerveux²⁷.

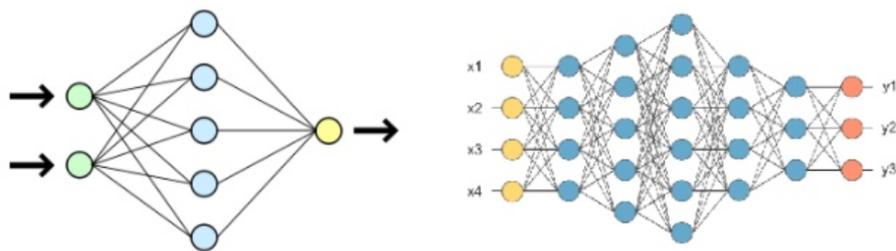
²⁷ Les neurones ont donc deux propriétés physiologiques : l'excitabilité, c'est-à-dire la capacité de répondre aux stimulations et de convertir celles-ci en impulsions nerveuses, et la conductivité, c'est-à-dire la capacité de transmettre les impulsions. Le cerveau n'est pas le seul organe du corps humain à comporter des neurones. Tel est également le cas de l'intestin, qui en comporte 200 fois moins, et du cœur.



Nicolas Rougier, schéma biologique d'un neurone, 2007

Pour apprendre, un cerveau humain ajuste la force des interconnexions entre ses milliards de neurones. Personne ne le programme. Un réseau de neurones artificiels fonctionne de la même manière²⁸. Il faut se le représenter sous les traits d'une série de variateurs de lumière reliés entre eux. Chacun reçoit des données en entrée et, suivant le réglage du variateur, produit une valeur en sortie. Cette opération se répète des milliers, voire des millions de fois à l'intérieur du réseau, chaque neurone pouvant envoyer son résultat à plusieurs autres. On parle d'apprentissage profond (*deep learning*) car le réseau est constitué de différentes couches de neurones qui communiquent. Au terme de l'analyse des données, les résultats des neurones de sortie sont associés afin de produire une valeur statistique. L'intelligence artificielle, par exemple, estime qu'il est à 85 % certain qu'une image *i* est celle d'un chat. Elle peut alors affirmer « il y a un chat dans *i* ». Un réseau de neurones est donc une chose simple à première vue : en fonction des données reçues en entrées, des connexions se déclenchent ou non entre les neurones, jusqu'à la sortie qui est le résultat du cheminement suivi. Et les paramètres internes s'ajustent progressivement en fonction des réussites et des erreurs.

²⁸ C'est en 1958 que le psychologue américain Frank Rosenblatt inventa la première machine à base de neurones artificiels, avec des entrées qui reçoivent les données, une couche interne qui effectue les calculs et une couche de sortie produisant des résultats. Il démontra qu'une telle machine pouvait potentiellement effectuer tous types de calculs. Mais, à l'époque, on était dans l'incapacité d'apprendre à de telles machines quoi que ce soit. C'est dans les années 1980, avec les travaux de Yann Le Cun, Yoshua Bengio et Geoffrey Hinton, qu'on parvint à inventer des techniques d'apprentissage permettant de faire progresser de telles machines.



Méthode:

1. Commencer par des valeurs choisies au hasard.
2. Faire varier légèrement les valeurs.
3. Soumettre en entrée les caractéristiques d'une situation bien connue.
4. Comparer le résultat des valeurs courantes avec celui des valeurs précédentes.
 - Si le nouveau résultat est plus proche de la réalité que le précédent, « récompenser » le paramétrage actuel et continuer dans cette direction.
 - Sinon, « affaiblir » le paramétrage actuel et chercher d'autres valeurs légèrement différentes qui pourraient donner de meilleurs résultats.
5. Recommencer en (3) jusqu'à ce que les résultats du réseau soient suffisamment proches des résultats attendus.

CETIC, juin 2017

Pour qu'un réseau de neurones soit efficace, il faut qu'il ait été supervisé à partir d'une base de données la plus vaste possible²⁹. À ce niveau, cerveau artificiel et cerveau naturel fonctionnent très différemment. Les tsunamis de données sont anti-biologiques : un cerveau humain est incapable de traiter autant d'informations mais est incommensurablement plus efficace pour comprendre les choses au départ de quelques données³⁰. Au départ, néanmoins, un réseau de neurones artificiels est totalement incapable de distinguer un chat d'un chien comme l'est le cerveau d'un bébé qui n'a encore rien appris. Il produit donc des résultats aléatoires, avec 50 % d'erreurs si son alternative est binaire. Le programme de supervision lui indique alors quels sont les cas dans lesquels il s'est trompé et ceux dans lesquels il a abouti au bon résultat. C'est à cet instant que les variateurs de lumière entrent en action. Le programme remonte le cours du réseau et fait varier l'intensité des neurones pour que le résultat de sortie se rapproche du résultat attendu. L'opération doit être

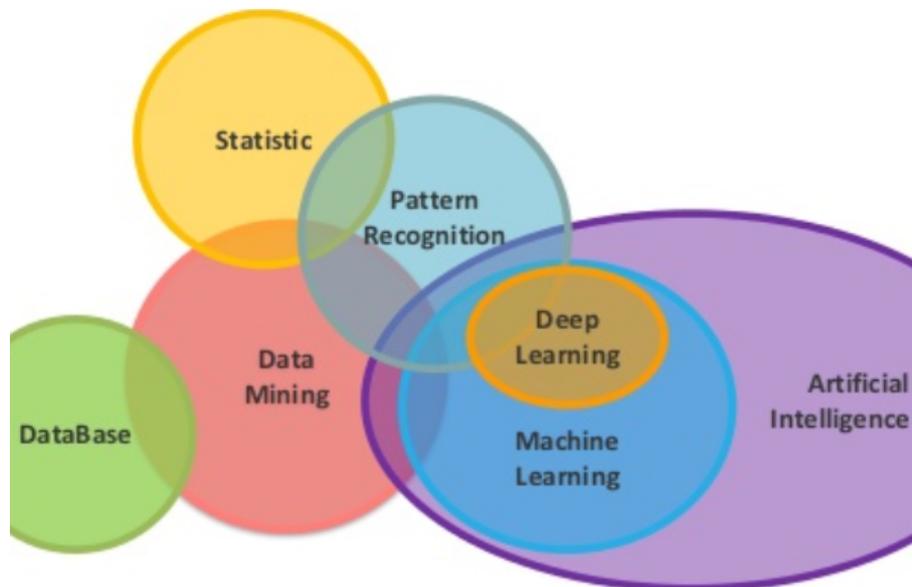
²⁹ Mais la quantité de données d'une seule image peut constituer une difficulté. Une photographie numérique contient plusieurs millions de pixels, chacun pouvant être considéré telle une donnée d'entrée pour le réseau. Le volume de données risque alors d'être beaucoup trop important pour être traité efficacement. C'est ici que l'aspect convolutif du réseau intervient — en anatomie, cela désigne les circonvolutions, les sinuosités du cerveau. Le réseau convolutif découpe l'image en plus petites parties et applique sur elles une série de filtres. Cela permet de réduire la quantité de données tout en conservant la pertinence des informations obtenues, notamment en identifiant des formes. La production de ces filtres dépend également du cycle d'apprentissage, c'est-à-dire que le réseau convolutif va créer progressivement des filtres optimisés pour les formes caractéristiques du sujet à reconnaître. Cette technologie peut aussi être déployée concernant du son ou d'autres types de données. Il suffit en effet de pouvoir représenter les informations sur une surface en deux dimensions (une image en quelque sorte) et de satisfaire à la contrainte principale des réseaux convolutifs : deux données proches ont de grandes chances d'avoir un rapport entre elles (comme deux pixels collés dans une image font généralement partie du même objet).

³⁰ L'information circule dans le cerveau, dans les neurones, à un mètre par seconde contre la vitesse de la lumière dans les microprocesseurs, soit 300 000 km par seconde.

reproduite un grand nombre de fois pour que le réseau, au fur et à mesure, ajuste ses résultats et fasse preuve de plus en plus d'acuité dans sa reconnaissance des images.

Par le passé, les systèmes d'analyse statistique exigeaient de fortes connaissances du problème et des experts œuvraient durant des mois à la transformation des données de la bonne manière pour que les algorithmes arrivent à en dégager des éléments pertinents et utiles. L'apprentissage profond est un grand progrès car il délègue cette partie de traitement des données au système d'apprentissage lui-même. Cette méthode peut donc être appliquée à des mégadonnées qui, par le passé, étaient laissées de côté à cause du coût de conception du système. L'intelligence artificielle peut ainsi être dix, cent, mille, un million voire un milliard de fois plus performante que les meilleurs cerveaux humains. Mais elle se limite à du *statistical learning*, de l'apprentissage statistique, inductif, permettant de dégager des règles générales à partir de cas particuliers. Par exemple, grâce à l'historique de nombreux acheteurs, un algorithme va déterminer que deux biens sont complémentaires parce que l'achat de l'un entraîne souvent, peu de temps après, l'achat de l'autre. Il pourra dès lors recommander automatiquement l'achat de l'un aux acheteurs de l'autre.

En résumé, l'intelligence artificielle a donc besoin de bases de données, de statistiques, de la science des données, de la reconnaissance de formes et de l'exploration et extraction de données. Et son cœur est constitué par l'apprentissage automatique, dont le propre cœur est l'apprentissage profond.



CETIC, juin 2017

L'intelligence artificielle apparaît sitôt qu'on parvient à obtenir de façon automatique, concernant un objet donné, des résultats pertinents. Un thermomètre est ainsi une forme très primaire d'intelligence artificielle car capable de saisir et exprimer le chaud et le froid, les variations de température, cela de façon plus précise qu'un homme dont le ressenti est biaisé par différents facteurs. De la plus sommaire à la plus perfectionnée, les intelligences artificielles sont donc susceptibles de connaître un très grand nombre de degrés d'évolution. Une IA consciente d'elle-même, donc une IA vivante, capable de développer ses propres projets, serait le quatrième âge de l'intelligence artificielle, après les algorithmes traditionnels (premier âge), l'apprentissage profond (deuxième âge) et l'IA contextuelle et conceptuelle, capable de créativité, d'empathie et d'arbitraire, de transversalité et de multidisciplinarité (troisième âge). Mais il est très incertain qu'on parvienne un jour à dépasser le deuxième âge. Pour l'heure, la machine peut être plus intelligente que l'homme, voire beaucoup plus intelligente que lui, mais seulement s'agissant de tâches très

spécifiques — comme elle peut être beaucoup plus forte physiquement que lui pour réaliser certains actes très précis et prédéfinis.

Autant les résultats produits grâce à des traitements statistiques sont impressionnants et peuvent intervenir dans une très grande variété de domaines, autant il est difficile de créer une intelligence artificielle fonctionnant autrement qu'à base de statistiques et d'exemples. Depuis 1956 et la conception de l'intelligence artificielle, très peu de progrès ont été faits en matière de vraie intelligence artificielle, c'est-à-dire de conscience artificielle, d'arbitraire artificiel, de compréhension artificielle, de créativité artificielle et de bon sens ou sens commun artificiel³¹.

La recherche fondamentale en intelligence artificielle se concentre désormais sur l'apprentissage faiblement supervisé ou non supervisé, c'est-à-dire l'apprentissage par exploration et par renforcement. Il s'agit de permettre aux machines d'apprendre avant tout par libre observation, sans leur fournir au préalable des données prémâchées, conquérant ainsi cette autonomie sans laquelle il n'est aucune forme d'intelligence. L'apprentissage non supervisé a lieu lorsque seulement de grands nombres d'exemples sont fournis à la technologie d'IA. À partir de ces exemples, cette dernière doit déterminer les structures et règles cachées permettant de relier les exemples entre eux et donc de comprendre ontologiquement ce dont il s'agit.

L'intelligence artificielle du jeu de go AlphaGo apprend ainsi les meilleures stratégies en jouant des millions de parties contre elle-même avec comme uniques connaissances de départ fournies par des humains les règles du jeu. Cette forme d'apprentissage par auto-entraînement est très différente des générations précédentes d'intelligence artificielle. Elle permet à l'IA d'être créative, d'aboutir à des solutions auxquelles l'homme n'aurait jamais pensé, mais elle requiert de grandes capacités informatiques puisqu'il faut l'équivalent de millions d'années de réflexion humaine pour obtenir des résultats pertinents. Aussi l'apprentissage automatique par exemples reste-t-il le plus courant et, dans de nombreux cas trop complexes pour permettre un apprentissage par exploration, le seul envisageable.

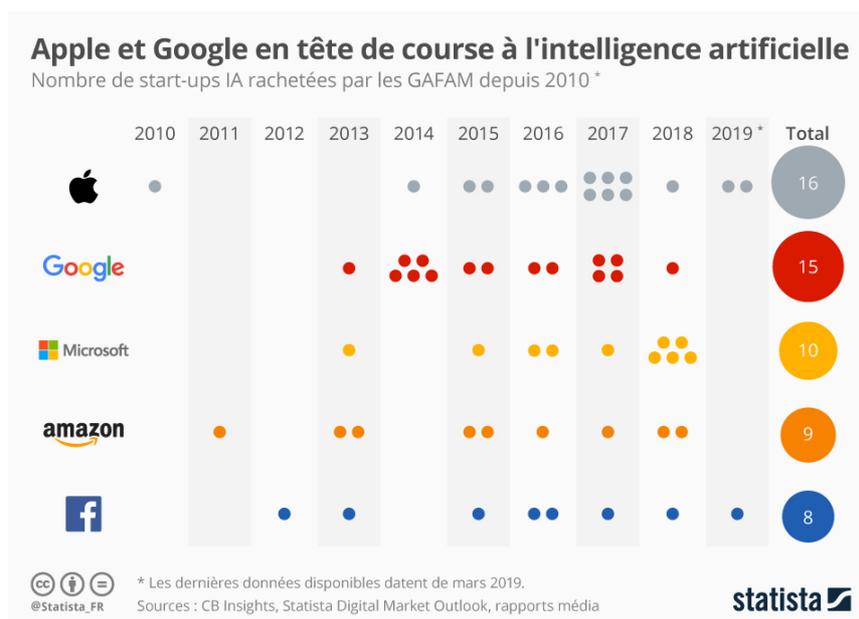
La recherche appliquée, de son côté, se charge d'optimiser, de rendre concrètes, de diffuser et d'exploiter les possibilités de l'apprentissage profond dans un maximum de domaines. C'est ainsi que l'intelligence artificielle s'est immiscée, assez discrètement et subrepticement d'ailleurs, au cœur de beaucoup d'objets et d'activités. En témoignent, en premier lieu, les instruments de communication numérique : ordinateurs, téléphones portables et tablettes numériques. Obtenir des images ou des informations très pertinentes à partir d'une requête donnée, grâce à des moteurs de recherche en ligne accessibles « en un clic », constitue par exemple une performance algorithmique extraordinaire dont le grand public profite quotidiennement, sans d'ailleurs s'en émerveiller outre mesure.

Cette recherche privée et appliquée peut être en avance par rapport à la recherche fondamentale. Néanmoins, s'il faut de nombreuses années, des décennies, pour résoudre un problème particulier relatif par exemple à l'apprentissage non supervisé, alors les travaux des centres de recherche publics resteront indispensables aux progrès de l'intelligence artificielle, les puissances privées risquant de refuser de financer cette recherche. Cela doit toutefois être relativisé par le fait que les multinationales du numérique embauchent de plus en plus de chercheurs, auxquels elles offrent des conditions de travail avantageuses, afin qu'ils réalisent leurs recherches fondamentales chez elles et pour elles — quoique les résultats de ces recherches sont souvent publiés en *open access* et *open source*.

La majorité des innovations numériques et des technologies et applications de l'intelligence artificielle provient d'un écosystème public-privé unique : la Silicon Valley. Cela explique pourquoi

³¹ En revanche, il est remarquable qu'une étude d'un cabinet de prospective américain a, en 2015, fait du métier de statisticien le numéro un dans le classement des plus beaux métiers du monde. Il s'agit un peu d'une revanche pour cette profession longtemps déconsidérée par les mathématiciens et, plus généralement, par les scientifiques. Dans ce genre de classement, le métier de mathématicien figure également souvent sur le podium.

d'aucuns qualifient la France et les pays d'Europe de « colonies numériques » ou de « pays du tiers-monde numérique », à défaut d'avoir vu naître certaines des grandes plateformes capables de créer et entraîner les IA³². Une forte concentration frappe d'ailleurs la production d'intelligences artificielles : il n'existe qu'une dizaine de grandes entreprises développant ces technologies (les fameux GAFAM et BATX : Google, Apple, Facebook, Amazon, Microsoft, Baidu, Alibaba, Tencent, Xiaomi³³). La concentration est d'autant plus forte que les GAFAM n'hésitent pas à racheter à prix d'or les start-up qui créent des IA performantes et/ou originales³⁴.



Reste que le fonctionnement de l'intelligence artificielle à travers des réseaux de neurones est délicat à cerner, même de façon abstraite et vague. On ressent forcément l'impression d'avoir affaire à quelque-chose d'insaisissable, d'incompréhensible, d'inmaîtrisable. Même pour un spécialiste, cet assemblage de millions de variateurs de lumière reliés entre eux, une fois qu'il en a terminé avec son apprentissage, s'avère relativement opaque. On ne saurait comprendre le rôle précis de chaque neurone artificiel dans la production d'un résultat ni retracer exactement la série de calculs effectuée par la machine. On ne peut que faire confiance au réseau, constatant qu'il produit généralement des résultats pertinents. C'est ce qu'on appelle communément la « boîte noire » de l'intelligence artificielle. Une fois que ses concepteurs, ses parents, lui ont donné la vie, elle atteint rapidement l'adolescence puis l'âge adulte et il leur devient difficile de lui imposer la conduite à adopter. Or elle est amenée à prendre des décisions immédiates et très nombreuses, comme en matière de conduite autonome, de reconnaissance de la parole ou des visages, de classement de fils d'informations, de recherche d'images, d'identification de contenus violents ou autrement choquants ou de détermination des publications à afficher en priorité sur les réseaux sociaux. De graves problèmes moraux peuvent se poser, par exemple en cas d'accident provoqué par une voiture autonome ; et il est difficile d'abandonner leur résolution aux calculs mathématiques et opaques d'un algorithme, au parcours incertain et hyper-complexe d'un réseau de neurones artificiels.

Les succès de l'intelligence artificielle dépendront de la confiance que le grand public lui portera ; et cette confiance suppose transparence et traçabilité — impératifs éthiques et politiques — dès lors qu'une décision est prise sur la base d'une indication algorithmique, cela afin d'éviter sa

³² France et Europe importent massivement les technologies IA américaines sur smartphones Android et iPhones.

³³ On peut y ajouter IBM, Samsung, Qualcomm, Intel, Philips, Siemens et Sony.

³⁴ L'IA est présentée comme un « grand aspirateur à valeur » et Laurent Alexandre, lors de son intervention au colloque du Sénat sur l'intelligence artificielle du 19 janvier 2017, cite l'exemple suivant : les 55 salariés de Whatsapp ont créé en quatre ans 23 milliards d'euros de valeur contre la moitié à peine pour les 130 000 ouvriers de Peugeot en un siècle.

non-explicabilité. Si une science en devenir est la rétro-ingénierie, consistant à déconstruire les raisonnements des machines, créer des technologies plus lisibles et traçables permettrait de mieux les comprendre et donc mieux les accepter. Mais il est très incertain que de telles technologies soient concevables en matière d'IA. Peut-être faudrait-il surtout que davantage de sociologues, d'anthropologues et de philosophes accompagnent les projets d'intelligences artificielles, de leurs prémices scientifiques à leurs conséquences pratiques ; et que les futurs ingénieurs et mathématiciens soient formés aux enjeux éthiques de leurs travaux, formés à la « robéthique ». Où revient le besoin de pluri-, inter- et même trans-disciplinarité dans les travaux relatifs à l'intelligence artificielle, besoin auquel ce livre tente de répondre en envisageant l'IA « dans toutes ses dimensions » — et notamment dans ses dimensions humaine, sociale, politique et économique.

III. Une problématique humaine, sociale, politique et économique : l'intelligence artificielle et les données, moteur et carburant d'un changement de civilisation

Déjà en 1954, dans son essai *La Technique ou l'enjeu du siècle*, alors que l'informatique commençait seulement à sortir de l'âge du simple concept, Jacques Ellul interrogeait le changement de nature de la technique dans la société, cela en des termes qui résonnent fortement à l'ère des IA et des multinationales du numérique. Auparavant simple instrument permettant à l'homme de mieux s'appuyer sur la nature et de dépasser sa condition initiale, elle muerait en un processus autonome de moins en moins au service des hommes car tendant à les enfermer dans de nouvelles formes de servitude. Le sociologue définissait « le phénomène technique [comme] la préoccupation de l'immense majorité des hommes de rechercher en toutes choses la méthode absolument la plus efficace »³⁵, ce qui présageait du mouvement d'hyper-rationalisation actuel de toutes les décisions et de toutes les actions. Et Jacques Ellul d'observer que la technique serait « à la fois sacrilège et sacrée » : « L'invasion technique désacralise le monde dans lequel l'homme est appelé à vivre. [...] Mais nous assistons à un étrange renversement : l'homme ne pouvant vivre sans sacré, il reporte son sens du sacré sur cela même qui a détruit tout ce qui en était l'objet : sur la technique »³⁶.

Aussi les puissantes technologies et les start-up qui les développent, dont certaines deviennent en un temps record des multinationales et même les premières puissances du monde contemporain, se retrouvent-elles statufiées, érigées en grands mythes des sociétés postmodernes. Beaucoup des services à base d'IA mis sur le marché ne sont pas des gadgets mais de véritables outils de gouvernement, qui régulent effectivement les comportements individuels et sociaux, parfois mieux que les lois des États.

Les grands acteurs actuels de l'économie, et des sociétés qui suivent cette économie, sont des puissances privées qui concurrencent les puissances publiques sur la scène interne et, plus encore, sur la scène externe — à tel point que le Danemark a cru bon de nommer un « ambassadeur auprès des GAFAM », ainsi explicitement considérés tels des acteurs souverains, tels les alter egos des États³⁷. Ces multinationales du numérique se sont développées autour d'algorithmes, sortes de recettes de cuisine jalousement conservées, les ingrédients étant les gigantesques masses de données constitutives du *big data*. Mieux vaut d'ailleurs avoir un mauvais algorithme avec beaucoup de données qu'un bon algorithme avec peu de données. Mais ces acteurs possèdent les deux.

³⁵ J. Ellul, *La Technique ou l'enjeu du siècle*, Armand Colin, coll. Sciences politiques, 1954, p. 18-19.

³⁶ *Ibid.*, p. 130-132.

³⁷ Et, lorsque cet ambassadeur souhaite rencontrer les représentants de Facebook, celui-ci dépêche ses stagiaires pour le recevoir.

L'intelligence artificielle devient ainsi le noyau dur de beaucoup de secteurs de l'économie, de la recherche et, par suite, de la société et de la vie. Ce sont les géants du numérique, GAFAM et BATX en tête³⁸, qui possèdent les principaux réservoirs de données. C'est pourquoi ils s'imposent telles les principales puissances technologiques et économiques, mais aussi comme les premières puissances politiques et scientifiques.

L'opinion publique s'insurge généralement haut et fort dès qu'un parlement ou un gouvernement a l'idée d'une loi sur le renseignement afin de lutter plus efficacement contre certaines menaces actuelles, un pareil texte étant perçu comme par trop liberticide. Mais, aujourd'hui, les véritables espions et les véritables services de renseignement sont ceux des multinationales des données. Ce sont elles qui connaissent dans leurs moindres détails tous les faits et gestes des individus. Alors que le « technolibéralisme » serait la grande mutation du libéralisme, devenant libéralisme augmenté³⁹, la véritable question des contre-pouvoirs est désormais celle des contre-pouvoirs face aux géants de l'économie numérique. Beaucoup acceptent sans autre forme de procès, car ils en sont trop peu conscients, la visibilisation intégrale de la vie sociale et même de la vie privée. Dans *1984* de George Orwell, le télécran⁴⁰ possède des angles-morts grâce auxquels on peut échapper à la surveillance. Aujourd'hui, ces moyens de se soustraire au contrôle sont le refus des réseaux sociaux et des objets connectés. Mais il est difficile de se convaincre d'un tel choix car le *Big Brother* omnipotent, précisément identifié et dont les missions étaient précisément définies, a laissé la place à une diversité de services *a priori* beaucoup moins effrayants et dont ce que l'on voit surtout est l'utilité — immense et incontestable dans d'innombrables domaines.

Deux dynamiques s'associent afin de dessiner le futur à bases d'algorithmes, aboutissant à redéfinir l'économie, la société, mais aussi la démocratie, les politiques publiques, et peut-être même l'homme en tant que tel. La première est la numérisation de la plupart des actes et activités individuels et collectifs, nourrissant de gigantesques bases d'informations plus ou moins personnelles et plus ou moins confidentielles⁴¹. Il y a peu encore, toutes ces données n'étaient pas enregistrées et, lorsqu'elles l'étaient, elles restaient difficiles d'accès et complexes à exploiter. En particulier grâce à l'internet (avec les sites web, les applications mobiles et les objets connectés), des torrents d'informations sont en permanence livrés par chacun et stockés dans d'immenses serveurs informatiques.

Quant à la seconde dynamique, elle est liée au besoin de comprendre et interpréter ces mégadonnées. C'est ici qu'interviennent les algorithmes, donnant aux ordinateurs des instructions mathématiques pour trier, traiter, agréger et représenter ces données. On assiste alors à un changement de statut des technologies de traitement de l'information : désormais, elles vont très au-delà des trois fonctions de base que sont la collecte, le stockage et l'indexation. De puissantes technologies statistiques et « prédictives » se développent en suivant l'augmentation des capacités de calcul des ordinateurs. L'intelligence artificielle est capable d'identifier des corrélations, de percevoir des ressemblances ou de trouver des informations clés parmi des ensembles gigantesques de données beaucoup mieux que l'intelligence humaine. Les classements, palmarès, compteurs, cartes, recommandations et notes diverses qui s'affichent en permanence sur les écrans des

³⁸ Mais d'autres acteurs sont bien sûr importants, à l'image, par exemple, d'Uber ou de Netflix.

³⁹ É. Sadin, *La silicolonisation du monde – L'irrésistible expansion du libéralisme numérique*, L'Échappée, coll. Pour en finir avec, 2016.

⁴⁰ Le télécran est à la fois un système de télévision, qui diffuse en permanence les messages de propagande du Parti, et un outil de vidéo-surveillance permettant à la Police de la Pensée d'entendre et de voir ce qui se fait dans chaque pièce où se trouve un individu.

⁴¹ En matière de données, la France, traditionnellement jacobine, est fort bien dotée. Le mouvement d'ouverture des données (*open data*) est d'autant plus bienvenu que beaucoup de jeux de données massives sont maîtrisés par des acteurs publics. Par exemple, en matière de santé, l'assurance maladie française est une ressource de données sans équivalent dans le monde.

ordinateurs, des tablettes et des téléphones portables sont les résultats de la calculabilité toujours plus fine des activités à venir en fonction des traces laissées par les activités passées.

La quantité de données produites par la société et les hommes numériques croît de façon exponentielle, tout comme les performances des réseaux de communication, des systèmes de stockage et des algorithmes. L'intelligence artificielle touche (ou touchera) tous les domaines : elle réaménage les villes, facilite l'accès au savoir et la diffusion de la culture, permet une médecine personnalisée et plus efficace, sert à lutter contre le changement climatique, tente de trouver des emplois pour les personnes au chômage, affecte les étudiants dans les établissements d'enseignement supérieur, décide d'accorder ou non des prêts bancaires, intervient dans les placements financiers, indique aux automobilistes les routes à emprunter, diffuse des publicités et autres contenus en rapport avec les besoins et/ou avec les centres d'intérêt, forme les couples etc. Au résultat, l'IA serait le moteur d'un changement de civilisation avec la marchandisation intégrale des vies et l'algorithmisation de pans entiers de l'évolution des sociétés⁴².

À l'image de toutes les grandes révolutions économiques et sociétales, amorcées par la découverte et l'exploitation d'une nouvelle forme d'énergie, les données numériques constituent une ressource, un gisement de valeur capable de relancer l'innovation, la productivité et la croissance tout en les guidant vers de nouveaux horizons jusqu'alors inconnus. Mais ces données peuvent aussi être utilisées afin de créer de nouvelles formes de servitude. Peut-être même s'agit-il d'une nouvelle forme de servitude volontaire : en cliquant systématiquement et machinalement sur « j'accepte les conditions d'utilisation et les cookies », car il est sinon impossible ou difficile d'utiliser beaucoup des services en cause, chacun consent plus ou moins consciemment à une extorsion de ses données.

C'est pourquoi la politique et le droit doivent intervenir⁴³, à l'image de l'Union européenne et du *Règlement général sur la protection des données* (RGPD) applicable depuis le 25 mai 2018, qui vise à mieux protéger les données des européens⁴⁴. Au-delà, se pose avec de plus en plus d'acuité la question d'un système dans lequel chacun serait maître de ses données, pourrait plus et mieux décider de celles qu'il abandonne, à qui et dans quelles conditions, ainsi que de celles qu'il conserve. En ce sens, certains, à l'image du philosophe Gaspard Koenig⁴⁵, défendent l'instauration d'un droit de propriété sur les données — quand celles-ci, pour l'heure, sont *res nullius*, des choses sans maître. Des start-up, anticipant de telles patrimonialisation, monétisation et commercialisation des données, travaillent déjà sur des portefeuilles intelligents de données qui permettraient de jouir pleinement de l'*usus*, du *fructus* et de l'*abusus* sur ses informations personnelles, *i.e.* les droits de

⁴² É. Sadin, *La vie algorithmique – Critique de la raison numérique*, L'Échappée, coll. Pour en finir avec, 2015.

⁴³ Il est d'ailleurs significatif que Mark Zuckerberg plaide pour plus de régulation publique car il est difficile pour des groupes tels que Facebook de produire les normes sur certains problèmes de société et certaines questions sensibles, par exemple afin de limiter la liberté de communication de certains mouvements radicaux, ce qui revient à les censurer. Dans tous les cas, il n'est sans doute pas réjouissant que des libertés fondamentales telles que, par exemple, la liberté d'expression ne reposent plus sur des sources publiques mais sur des sources privées, en l'occurrence les GAFAM dont les conditions d'utilisation peuvent être plus effectives que le premier amendement de la Constitution aux États-Unis ou la grande loi du 29 juillet 1881 en France.

⁴⁴ Le double objectif de ce texte est d'accroître la responsabilité des acteurs de traitements de données et de consacrer de nouveaux droits en faveur des citoyens européens : droit à l'effacement (ou droit à l'oubli), droit à la portabilité des données personnelles, consentement « explicite » et « positif » à la collecte de ses données personnelles (notamment via l'acceptation des cookies et le contrôle de l'utilisation des données renseignées à travers les formulaires de contact), droit de ne pas faire l'objet d'une décision fondée exclusivement sur un traitement automatisé, y compris le profilage, produisant des effets juridiques concernant la personne ou l'affectant de façon similaire, ou encore obligation pour toutes les activités qui peuvent avoir des conséquences importantes en matière de protection de données personnelles d'être précédées d'une étude d'impact sur la vie privée.

⁴⁵ G. Koenig, *La fin de l'individu – Voyage d'un philosophe au pays de l'intelligence artificielle*, Éditions de L'Observatoire, 2019. L'auteur rappelle combien, historiquement, toutes les grandes révolutions techniques ont été accompagnées d'une extension du droit de propriété, par exemple avec la terre et la paysannerie au sortir de la féodalité, l'imprimerie et le droit d'auteur qui l'a suivie, soutenu par Beaumarchais, ou encore la création des brevets sur les innovations au moment de la révolution industrielle.

les utiliser, d'en percevoir les fruits, les revenus (par exemple en les louant), et d'en disposer, de les transmettre ou de les détruire. Ce ne seraient dès lors plus les services mais les utilisateurs qui imposeraient leurs « conditions d'utilisation », acceptant d'être « nudgés » par des IA dans certaines de leurs activités et certains de leurs choix mais pas dans d'autres.

L'intelligence artificielle se construit aussi par la politique et par le droit, qui fixent des bornes, un cadre, afin que le progrès profite au plus grand nombre et non seulement aux plus forts. L'enjeu, de leur point de vue, est de parvenir à devancer plutôt que suivre les technologies et les usages. Dans le cas du RGPD et des régimes de protection des données personnelles plus généralement, ceux-ci sont un obstacle de taille aux progrès de l'intelligence artificielle dans les pays concernés puisque d'autres, ailleurs dans le monde, peuvent beaucoup plus librement exploiter de grandes bases de données, devenant ainsi les leaders de l'IA. L'Europe, en édictant le RGPD, a donné un avantage énorme aux géants du numérique américains et chinois. Mais est-ce critiquable ? Il s'agit d'un choix politique. Gaspard Koenig y voit le « suicide stoïcien de l'Europe », qui préfère mourir en brandissant ses valeurs et principes plutôt que de continuer à vivre en les bafouant⁴⁶. L'Europe préfère sanctifier certains de ses principes, notamment des droits et libertés fondamentaux, même si cela doit nuire à sa compétitivité, à sa capacité d'innovation et, finalement, à sa place sur l'échiquier international.

Lorsque la confidentialité des données relève de la protection de la vie privée en Europe, conçue tel un droit à prétention universelle, il apparaît suspect de vouloir protéger sa vie privée en Chine. Là-bas, chercher à protéger ses données est perçu comme de l'égoïsme de la part de quelqu'un qui ne contribue pas à l'effort collectif en vue du bien commun. De même, la Chine tire un grand avantage de la société de surveillance qu'elle déploie à grand renfort d'IA⁴⁷. Et, par exemple, elle profite de l'absence de liberté d'expression sur son territoire et de la consécration de cette même liberté ailleurs dans le monde, qu'elle exploite afin de diffuser sa propagande et sa désinformation et ainsi accroître son *soft power*⁴⁸. Mais faudrait-il pour autant suivre cet exemple ? Depuis des siècles, la prospérité allait de pair avec le développement et la protection des droits et libertés individuels. Avec l'économie des données et des IA, le phénomène inverse se produit et, plus généralement, divers droits et libertés fondamentaux se trouvent malmenés. Sont atteints ou pourraient être atteints le fonctionnement du système constitutionnel, la mécanique électorale, l'indépendance de la justice, la transparence de la vie publique et de l'administration, l'égalité et l'absence de discriminations, la protection de la vie privée et des données, la liberté d'expression, la liberté académique, la liberté de croyance et notamment de religion, le droit à un traitement équitable, les droits des personnes issues des minorités, ainsi que divers droits économiques et sociaux.

Cela pose d'innombrables dilemmes moraux, notamment du côté des États-Unis. Reste que l'objectif des multinationales de l'IA et des données est d'inviter leurs utilisateurs à leur livrer le maximum d'informations personnelles les concernant, afin de pouvoir les cibler le plus finement possible, ainsi que de générer par des « likes » ou « dislikes » de l'engagement, car c'est cela qui génère de la *data*.

⁴⁶ G. Koenig, *La fin de l'individu – Voyage d'un philosophe au pays de l'intelligence artificielle*, Éditions de L'Observatoire, 2019.

⁴⁷ Le témoignage le plus extrême de cette société de surveillance est peut-être celui-ci : à Pékin et dans d'autres villes chinoises, les distributeurs de papier de certaines toilettes publiques sont équipés de technologies de reconnaissance faciale afin de lutter contre les abus ; ainsi une même personne ne peut-elle utiliser plus de soixante centimètres de papier toilette dans un délai de neuf minutes — cette mesure est motivée par des considérations écologiques, donc louables, et il est en l'occurrence difficile d'avancer que d'autres motivations, moins légitimes, se cacheraient derrière.

⁴⁸ Dans le « Document 9 » de juillet 2012, un manifeste secret destiné aux cadres du parti communiste chinois qui prohibe expressément les analyses et opinions contraires à la ligne officielle, ce qui confine à une froide interdiction de toute opposition, y compris médiatique, la liberté de la presse est classée parmi les « sept périls occidentaux », aux côtés notamment des droits de l'homme et de l'indépendance de la justice.

Malgré les différences fortes qui existent d'un point de vue juridique, éthique et idéologique, la Chine ne souhaite absolument pas se couper du monde et, tandis que certains GAFAM s'établissent sur son sol afin de profiter de certains de ses talents, la *Digital Silk Road* (« route de la soie numérique ») vise à connecter l'Union Européenne avec la Chine à l'aide de différents types d'infrastructures, dont les satellites, la 5G et les câbles sous-marins. Ce projet pourrait permettre à la Chine, d'ici 2030, de devenir le « principal centre d'innovation en intelligence artificielle du monde, transformant le pays en chef de file mondial de l'innovation et faisant de lui la plus grande puissance économique », suivant les termes du *Plan national de la Chine pour l'IA*. Surtout, le gouvernement et la population chinois sont mieux préparés qu'ailleurs à la révolution des intelligences artificielles, parce qu'ils y voient le principal levier économique afin de devenir la première puissance mondiale et parce que ces technologies profitent généralement aux collectivités plus qu'aux individualités. Elles sont donc perçues comme collectivistes et utilitaristes, ce qui est en parfait accord avec la philosophie confucéenne — qui connaît d'ailleurs un fort renouveau en Chine.

Rien n'est plus égoïste qu'un État. Historiquement, un pays qui domine les autres technologiquement et économiquement a toujours été impérialiste. Même si la Chine annonce travailler aux progrès de l'humanité⁴⁹, il est difficile de croire qu'elle n'ait pas de visées impérialistes. Le but d'une guerre n'est pas de conquérir des territoires mais de conquérir des populations. Longtemps, cela passait par l'invasion du territoire, mais, aujourd'hui, il devient possible de conquérir des populations à distance, grâce au *soft power*, en s'imposant dans les esprits et dans les pensées. Tout en demeurant dans la campagne de sa Franche-Comté natale, on peut devenir « sujet » de l'empire chinois en lui livrant ses données et en utilisant ses services.

Par ailleurs, les conséquences de l'IA qui préoccupent le plus ses observateurs et commentateurs sont peut-être celles sur le travail. Les tâches nouvellement attribuées à l'informatique étaient pour la plupart dévolues à des humains, qui perdent en conséquence leurs emplois. Et rien n'assure qu'une redistribution efficace soit possible. Énormément de nouveaux métiers vont être créés, mais rien n'indique que les actifs d'aujourd'hui et de demain seront en capacité d'occuper ces postes très spécifiques tels que *data analyst* ou *data scientist*⁵⁰. *Data scientist* serait pourtant le « métier le plus sexy du XXIe siècle », selon la *Harvard Business Review*. Mais le World Economic Forum s'inquiète du fait que les travailleurs, notamment français, seraient très mal armés pour s'adapter et affronter le changement. L'intelligence artificielle pose donc, à moyen terme et même à court terme, la question d'une aggravation de la situation du marché de l'emploi.

Dans le pire des scénarios, l'IA pourrait ne pas générer un grand remplacement des métiers et produire une masse de gens ne vivant, par exemple, que grâce au versement d'un revenu universel⁵¹. Afin de demeurer compétitif, à l'échelle de l'individu et à l'échelle de la société, dans un monde où l'IA, même faible, pourra effectuer toutes les tâches sans valeur humaine ajoutée (conduire un camion, tenir une comptabilité, traduire des textes ou des discours, diagnostiquer une maladie, par exemple), d'importants moyens doivent être consacrés au renouvellement et à l'amélioration de l'enseignement et de la recherche. Dans le cas contraire, le risque est celui d'un « apartheid intellectuel »⁵², l'intelligence artificielle ne permettant plus qu'à une frange de la population de travailler, parce qu'elle resterait dotée d'une intelligence supérieure et, par suite, d'une utilité économique et sociale.

⁴⁹ Cf. Kai-Fu Lee, *I.A. La plus grande mutation de l'histoire*, Les Arènes, 2019.

⁵⁰ Le *data analyst* et le *data scientist* doivent : traduire un problème business en problème mathématiques et statistiques ; identifier les sources de données pertinentes ; concevoir des « entrepôts de données » ; évaluer les données, les traiter et les resituer dans le système d'information cible. Ces professionnels combinent une triple compétence : expertise statistique et informatique, connaissance des bases de données et de l'informatique, expérience métier dans leur secteur d'activité.

⁵¹ Il s'agit par conséquent d'une idée très libérale en ce qu'elle prend acte de l'impossibilité de trouver un emploi pour chacun, donc acte de la fin du travail.

⁵² L. Alexandre, *La guerre des intelligences – Intelligence artificielle versus intelligence humaine*, JC Lattès, 2017.

S'il n'y a plus besoin d'hommes pour les métiers basiques, parce que l'IA s'en charge plus efficacement et à un moindre coût, plus rapidement et avec moins d'erreurs et d'approximations — il faut 20 ans pour « fabriquer » un comptable et 30 ans pour « fabriquer » un médecin, mais seulement quelques jours pour fabriquer un algorithme aussi performant que ce comptable et que ce médecin —, et si ne comptent plus que les capacités cognitives complémentaires de l'IA, alors il faudra non seulement inventer des métiers à haute valeur humaine ajoutée, mais aussi préparer les individus afin qu'ils puissent exercer ces métiers, former non des techniciens et des applicateurs mais des tacticiens et des créateurs, entre autres. L'homme devra se spécialiser dans les secteurs requérant des compétences multidisciplinaires et transversales⁵³, où il y a peu de données et de statistiques mais un fort besoin de sens commun, de souplesse, d'adaptabilité, de créativité, de culture générale, de sens critique, d'empathie, d'affectivité, d'humanité, de faculté à innover, de capacité à prendre des décisions arbitraires, ainsi que, parfois, de force physique et de dextérité. L'IA est gratuite en valeur relative par rapport à l'intelligence biologique⁵⁴, ce qui emporte des conséquences fortes sur le marché de l'emploi car, lorsqu'un bien est gratuit, ses substituts ne peuvent plus le concurrencer et disparaissent, tandis que les complémentaires deviennent beaucoup plus recherchés et gagnent donc en valeur. Ici, le travail complémentaire est le travail qualifié sous l'angle de la valeur humaine ajoutée.

La formation professionnelle pourrait jouer un rôle essentiel, si elle devient plus efficace, mieux adaptée, plus visible et mieux valorisée. Plus généralement, les parcours scolaires, éducatifs et universitaires ne doivent pas laisser passer le train de l'IA. Il est ainsi remarquable que le nombre d'étudiants inscrits au cours introductif à l'apprentissage automatique dans cinq prestigieuses universités américaines⁵⁵ a en moyenne plus que quadruplé depuis 2012.

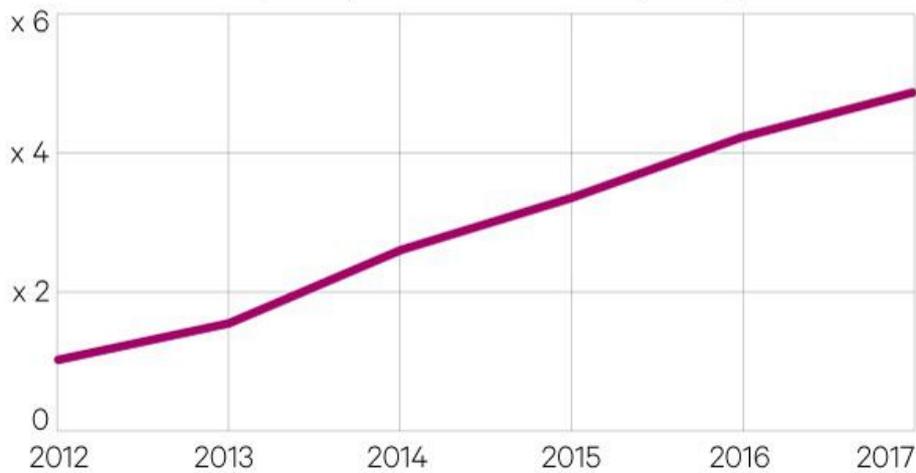
⁵³ L'IA peut concurrencer un radiologue mais pas un médecin généraliste. Face à l'IA, le généraliste garde toute sa place, au contraire du spécialiste.

⁵⁴ Mais il faut distinguer le prix des IA simplement informatiques, du secteur tertiaire, et celui des IA avec robotique en raison des coûts matériels et de fabrication souvent très élevés des robots. Le travail humain est beaucoup plus concurrencé par les premières que par les secondes. L'intelligence artificielle sert néanmoins le développement qualitatif et quantitatif de la robotique. Des robots thérapeutiques interviennent déjà dans certaines maisons de retraite. Des robots ménagers sont, quant à eux, présents depuis longtemps dans les foyers pour aider au quotidien. Des robots d'aide au déplacement permettent de pallier à certaines déficiences physiques. Et des robots-compagnons, sécurisants et divertissants, arrivent sur le marché, tandis qu'on devrait bientôt voir apparaître des robots majordomes au service des plus fortunés. Il faut aussi ajouter tous les robots non humanoïdes, tels que les véhicules autonomes ou les drones de livraison.

⁵⁵ Berkeley, University of Illinois Urbana-Champaign, Stanford, Carnegie Mellon et University of Washington à Seattle.

Inscriptions aux cours introductifs d'apprentissage machine aux Etats-Unis

Evolution des inscriptions par année universitaire, en moyenne



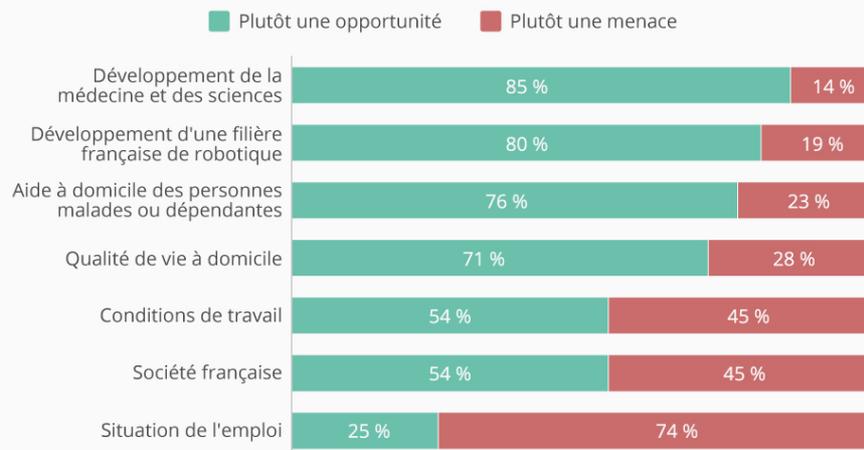
* LES ÉCHOS* / SOURCES : AI INDEX 2018, BERKELEY, CMU, STANFORD, UIUC, UW

La question de l'emploi est l'une des principales clés de l'acceptation sociale de l'intelligence artificielle. Les révolutions industrielles ont toujours généré un moment d'angoisse généralisée. L'incertitude liée à la nouveauté et au brouillard entourant l'avenir, en premier lieu à un niveau micro-économique et à l'échelle de l'individu-salarié, provoque des inquiétudes. Mais les nouvelles technologies, jusqu'à présent, ont toujours bouleversé les modes de production sans pour autant entraîner des pertes d'emplois supérieures aux créations d'emplois. De nouveaux secteurs d'activité vont donner lieu à des métiers auxquels l'on n'a pas encore idée. Reste que l'adaptabilité et la souplesse, dans un tel contexte, seront les conditions d'une transition la moins chaotique possible vers l'économie des données et des IA.

Ce sont donc logiquement les conséquences du déploiement des IA sur la situation de l'emploi qui suscitent le plus de craintes, tandis que, à l'inverse, ses répercussions en matière de développement de la médecine et des sciences génèrent le plus d'espoirs :

L'intelligence artificielle, chance ou menace ?

Avis des Français sur le développement de la robotique dans plusieurs domaines *



CC BY ND
@Statista_FR

* Les points de pourcentage manquants correspondent aux "ne sait pas".
Source : Odoxa

statista

Si, dans l'ensemble, l'optimisme semble l'emporter sur le pessimisme en matière de perspectives de l'intelligence artificielle, il est remarquable que le deuxième domaine au sujet duquel les craintes du grand public sont les plus fortes est la « société » en général. L'IA n'est ainsi pas le sujet intéressant le moins la « collapsologie », ce « discours sur ce qui s'effondre », en l'occurrence la civilisation du XIXe et du XXe siècles et le « système global » en raison du réchauffement climatique et de l'épuisement des ressources. Elle concerne aussi les survivalistes, qui brandissent les menaces économique, climatique, civilisationnelle et technologique et qui se préparent à une très prochaine fin du monde — ou fin d'un monde. Les collapsologues rappellent que plus une société est inégalitaire, plus elle risque d'exploser. L'IA pourrait par conséquent être le terreau d'une très grave crise sociale. En tout cas pose-t-elle de redoutables questions éducatives, philosophiques et politiques.

La démocratie pourrait ne pas survivre au creusement de fossés au sein de la population, notamment en termes d'intelligence, de culture et d'éducation. La « trumplexisation » de la fin des années 2010 en serait un premier témoignage. La montée de la misère, y compris misère intellectuelle, s'accompagne toujours d'une poussée du populisme et des partis extrémistes. Or le danger parmi les plus ultimes de tous, mis en lumière par l'historien israélien Yuval Harari⁵⁶, est celui de faire advenir un monde scindé en deux, avec des « hommes-dieux » d'un côté et des inutiles de l'autre, un monde gouverné par une petite aristocratie de puissants grâce à leur capacité d'innovation et à leurs moyens technologiques et un monde hostile pour les hordes de désœuvrés du numérique incapables de s'adapter aux changements de leur environnement. Le risque est d'arriver à une situation non plus de complémentarité entre les individus (intellectuels et manuels, directeurs et salariés, décideurs et exécutants) mais d'opposition entre une élite de l'intelligence et un prolétariat affaibli, démuni, abandonné, entre ceux qui voguent sur la vague des nouvelles technologies et les naufragés du numérique et de la société de la connaissance.

En outre, la démocratie est d'ores et déjà victime de la désintermédiation généralisée promue par les grands acteurs de l'économie des données et des IA, qui a pour effet de renforcer la défiance envers les représentants, envers les élus⁵⁷. Et bien d'autres défis et problématiques se posent : atteintes à la vie privée, pollutions, hiérarchisation arbitraire et non journalistique de l'information, incompréhension du grand public qui ne prend pas garde à l'enfermement dans des bulles de filtre et à la mise en pilotage automatique des vies, biais dans les données qui se reflètent dans les résultats, donc discriminations en tous genres — l'intelligence de l'intelligence artificielle dépend assez exactement de la qualité des données qui la nourrissent et une base de données biaisée et discriminante produira des résultats symétriquement biaisés et discriminants ; l'intelligence artificielle est sous l'influence des catégorisations, préconçus et stéréotypes de ceux qui l'abreuvent en données, si bien que les problèmes peuvent venir moins des algorithmes et des réseaux de neurones que des masses d'informations non neutres qu'ils traitent⁵⁸.

Cependant, si l'IA contribue à l'effondrement d'un monde qu'on pouvait croire stable et pérenne, elle participe aussi à la construction d'un nouveau monde — et peut-être cet aspect, constructif et positif plutôt que plaintif et négatif, est-il le plus intéressant. Il ne faut en tout cas pas exagérer l'un au détriment de l'autre. Il convient de n'être ni technophobe ni technophile mais « techno-lucide » et « techno-éclairé ». L'IA et les données pourraient être le prochain « cliquet malthusien », permettant à l'humanité de faire un grand bond en avant aussi bien d'un point de vue

⁵⁶ Y. Harari, *Homo Deus – Une brève histoire de l'avenir*, Albin Michel, 2017.

⁵⁷ De l'autre côté de la balance, l'objectif de ces acteurs est de garantir ou renforcer la confiance des utilisateurs dans leurs services, dont dépendent très largement leurs succès.

⁵⁸ C'est ainsi que l'on vit dans un monde où l'IA a envoyé des gens en prison parce qu'ils étaient noirs. De mêmes discriminations existent dans les recrutements. Par exemple, Amazon a développé un algorithme censé permettre de choisir les meilleurs candidats dans l'informatique et le candidat idéal s'est avéré être... un homme. On aboutit ainsi à des outils de recrutement qui disqualifient automatiquement toutes les candidatures féminines.

quantitatif que d'un point de vue qualitatif, après l'agriculture, la révolution industrielle, l'exploitation du charbon, du gaz, de l'électricité ou du pétrole.

Dans son rapport « Donner un sens à l'intelligence artificielle », rendu public le 28 mars 2018, Cédric Villani pose la question des ministères qui sont ou seront impactés par les développements de l'IA et auxquels, par conséquent, son texte devait s'adresser en priorité. Sa conclusion est que tous les ministères sont concernés : économie, travail, éducation, enseignement supérieur, recherche et innovation, santé, environnement, mais aussi justice, culture, armée, police, transports, action et comptes publics, aménagement du territoire, ville et logement. Il est vrai que la plupart des basculements civilisationnels en cours sont liés, au moins indirectement, à l'IA : de l'Atlantique vers la zone Asie-Pacifique (de Shangaï jusqu'à la Silicon Valley) ; du neurone vers le transistor ; de la loi du Parlement vers la loi des algorithmes ; des États-nations vers les multinationales de l'économie numérique ; d'une économie de subsistance à une économie de « sursistance » (tuer la mort⁵⁹, coloniser mars, augmenter le cerveau et l'intelligence humaine) etc.

L'intelligence artificielle n'est pas une simple innovation permettant de répondre autrement à un problème ponctuel et ancien. Elle constitue une véritable révolution en raison de sa transversalité — comme l'ont été les inventions du langage, de l'écriture, de l'imprimerie, de la télégraphie sans fil, de l'électricité ou de l'aviation — ; elle est le facteur clé d'autant de nouvelles solutions à explorer que de nouvelles difficultés à résoudre et de nouveaux défis à affronter. C'est pour cela que l'intelligence artificielle suscite à la fois chez certains un enthousiasme fanatique et chez d'autres une inquiétude horrifique. Mais cet enthousiasme et cette inquiétude demeurent relativement vagues et indéfinis, car il est aujourd'hui encore difficile de deviner dans quelles directions et en quels domaines les changements provoqués par les IA seront les plus importants et les plus radicaux.

Les plus grandes surprises proviennent d'ailleurs moins des nouvelles technologies que des usages sociaux ou économiques qui en sont faits. Les hommes fabriquent des machines qui, en retour, les reconstruisent. L'humanité sort forcément modifiée de la rencontre avec la technologie qu'elle a créée. De moins en moins de gestes du quotidien, d'achats, de déplacements, de décisions personnelles ou professionnelles ne sont pas dictées, plus ou moins intégralement, par des calculs et des statistiques. Avec un effet performatif fort, des IA aident à prendre des décisions ; et, souvent, elles les prennent elles-mêmes. Comme les inventions du microscope ou du télescope ont permis de voir autrement le monde et l'univers, les IA et les capteurs numériques, toujours plus omniprésents, mesurant tout et prédisant tout, redéfinissent le regard de l'homme sur son monde et sur lui-même.

Cela concerne y compris, et d'ailleurs dans une large mesure, la recherche et l'enseignement, donc la formation et la transmission des pensées et des connaissances, ainsi que de l'esprit critique, du sens moral, de la capacité de discernement et de doute. Les chercheurs en sociologie, économie, philosophie, histoire, droit etc. doivent se saisir de ces révolutions qui dictent et dicteront l'avenir de leurs disciplines, du point de vue des objets d'étude comme du point de vue des moyens d'étude. Et, parce que l'IA concerne tout le monde de très près, ils doivent se mettre au service du grand public et ne pas se contenter de travaux inter-universitaires, en vase clos. C'est pourquoi ce livre, par exemple, a été conçu afin d'être accessible à tous et donc destiné à tous.

⁵⁹ L. Alexandre, *La mort de la mort – Comment la technomédecine va bouleverser l'humanité*, JC Lattès, 2011. Pour Google et sa filiale Calico, il s'agit d' « euthanasier la mort ». La première personne qui vivra mille ans est peut-être déjà née, scandent certains transhumanistes.

IV. Une problématique scientifique : le besoin d'une science grand public de l'intelligence artificielle

Les contributions contenues dans cet ouvrage, rédigées par les meilleurs spécialistes de chaque question avec l'intention qu'elles soient les plus claires et synthétiques possible, interrogent l'intelligence artificielle dans toutes ses dimensions — à savoir sous un angle historique, sous un angle sociologique, d'un point de vue anthropologique, d'un point de vue économique, mais aussi sous un angle politique, en termes juridiques ou encore avec un regard socio-scientifique, philosophique ou prospectif. Par ailleurs, comme y invitent la collection Europe & Asie et les liens entre l'Université Paris 1 Panthéon-Sorbonne et l'Université Jiao Tong de Shanghai qui en sont à l'origine, quelques textes concernent la situation de l'intelligence artificielle sur le continent asiatique et plus particulièrement en Chine — pays qui constitue l'autre grand pôle de développement de ces technologies avec les États-Unis⁶⁰.

Forcément non exhaustif et ignorant de ce que sera le futur, l'ouvrage collectif *L'intelligence artificielle – Dans toutes ses dimensions* en appelle beaucoup d'autres. On est seulement à l'aube de l'« ère de l'intelligence artificielle ». Cela rend le sujet passionnant et les recherches en la matière indispensables et, pour longtemps, inépuisables. En même temps, cela voue ces travaux à une obsolescence rapide. Tel est le lot de toutes les problématiques qui obligent à envisager un futur aux multiples inconnues et innombrables possibilités plus qu'à se rappeler un passé déjà fait ou à observer un présent en train de se faire. Il n'en faut pas moins accepter d'interroger l'avenir, ne pas s'y fermer ni chercher à transposer artificiellement les cadres du passé et du présent dans le futur — ou alors on risquerait de commettre un « crime contre l'avenir ». Si le sujet est plein de mystère, du point de vue de l'évolution de la discipline, de celui des ressorts scientifiques profonds des technologies ou de celui de leurs conséquences dans tous les domaines, les chercheurs et les penseurs ne doivent pas s'en effrayer. Il s'agit pour eux d'une chance immense, d'une opportunité formidable, car ils pénètrent dans un vaste champ encore inexploré.

Au-delà du vecteur technologique, directement observable et objet de nombreux commentaires et approfondissements scientifiques, la profondeur et les méandres de ses intrications et implications sociales méritent d'être explorés. Les innovations de la technosphère mettent toujours un temps plus ou moins long pour infiltrer et conquérir la sphère symbolique et y imposer leurs lois iconoclastes. Ce n'est qu'en de nombreuses décennies, voire des siècles, que l'homme moyenâgeux est devenu pleinement « gutenbergien », moderne. De combien de temps l'IA aura-t-elle besoin pour transformer l'homme moderne en homme postmoderne ? Et le pourra-t-elle seulement ? De façon générale, le regard porté sur les technologies reste par nature hypermétrope : on voit mieux de loin que de près, mieux l'imprimerie que la télévision, mieux la télévision qu'internet, et mieux internet que l'intelligence artificielle.

Une radiographie critique de l'intelligence artificielle et de ses conséquences n'en est pas moins un enjeu démocratique et social fort. Dans tous les secteurs de la connaissance, les chercheurs doivent s'efforcer à comprendre les dynamiques à l'œuvre, encourager les progrès de tous ordres qui peuvent en être tirés et prévenir les dérives dont l'intelligence artificielle risque de

⁶⁰ Pour ne prendre qu'un exemple, le projet de « crédit social » du gouvernement chinois vise à mettre en place un système national de réputation des citoyens. Chacun se verrait attribuer une note, son « crédit social », fondée sur les grandes masses de données collectées par le gouvernement à propos de leurs activités. Ce système s'appuierait donc sur des outils de surveillance de masse et sur des algorithmes permettant d'analyser les innombrables données collectées.

constituer le terreau. Puisque tout le monde est concerné et impacté, tout le monde devrait pouvoir comprendre le fonctionnement de l'intelligence artificielle et identifier les moments lors desquels on agit sous l'influence d'une proposition algorithmique. Pourtant, habités par un sentiment d'incompétence et d'éloignement, bien que les IA soient partout et tout proches, on préfère souvent se désintéresser de la question, considérant qu'elle ne serait qu'affaire de spécialistes. C'est pourquoi l'intention de ce livre est d'offrir à tous un moyen de mieux cerner l'intelligence artificielle et son impact sur la société, sur la culture, sur l'économie, sur la vie.

Parmi les plus concernés, on trouve les médiologues. La médiologie est l'étude et la pensée des effets des médias et de tous les instruments relationnels. Cette discipline ne vise pas spécifiquement les médias compris comme presse, télévision, radio, web etc. « Médias » désigne ici ce qui se tient entre les personnes et permet de les relier, de les organiser. La médiologie s'intéresse donc aux milieux, cadres et guides, indissociablement sociaux et techniques, qui modèlent et remodelent les représentations symboliques des hommes et leur permettent de tenir ensemble, de former un « nous »⁶¹. La médiologie souligne combien « exister, c'est être relié. Aucun organisme ne peut se développer durablement à l'écart des autres, au point qu'un réseau de bonnes relations semble la condition *sine qua non* de nos vies »⁶². Les médiologues ont dès lors forcément beaucoup à dire et à écrire au sujet des IA, nouveaux vecteurs des relations sociales et nouveau fil conducteur des croyances et comportements collectifs ou groupusculaires. L'influence de ces technologies sur la richesse des jeux de communication, actuels ou futurs, est forcément immense et même, dans une large mesure, révolutionnaire. Ne devra-t-on pas bientôt reconsidérer entièrement l'écologie de l'esprit et de la pensée (l'étude des milieux dans lesquels l'esprit et la pensée se déploient, ainsi que de leurs rapports avec ces milieux⁶³) ?

Pourtant, l'outillage technique de la communication est régulièrement négligé dans l'expertise des phénomènes symboliques et informationnels. Or, avec l'IA, l'influence de cet outillage sur ces phénomènes devient encore plus déterminante qu'auparavant. La forme et le contenu d'une culture et de la société qui repose sur elle dépendent étroitement de matériaux techno-médiatiques et, par suite, d'un inconscient techno-médiatique, *a fortiori* à l'ère de l'intelligence artificielle. On a toutefois tendance à privilégier les données culturelles et symboliques et à délaisser les outils, infrastructures et réseaux techniques qui les portent et les meuvent. On connaît mieux l'histoire de la littérature que l'histoire de la librairie ou de l'imprimerie, mieux les œuvres d'art que les métiers, musées et matériaux qui permettent de les montrer au public. Les contenus sont pourtant étroitement dépendants des contenants. L'impact des technologies d'intelligence artificielle sur les modes de pensée est tel qu'il devient impossible de continuer à aborder les phénomènes en cause en observant leurs sommets visibles et en ignorant les bases immergées sur lesquelles ils reposent. Ces dernières sortent cependant de plus en plus de l'obscurité pour se révéler au grand jour.

L'intuition de Marshall McLuhan se confirme : « *The medium is the message* »⁶⁴ — *i.e.* la nature du média (le canal de transmission) l'emporte sur le contenu du message et, finalement, redéfinit son sens et sa portée. Ainsi, selon le philosophe canadien, qui écrivait dans les années 1960, « en réalité et en pratique, le vrai message est le médium lui-même, c'est-à-dire que les effets d'un médium sur l'individu ou sur la société dépendent du changement d'échelle que produit chaque nouvelle technologie, chaque prolongement de nous-mêmes, dans notre vie. Toutes les technologies créent petit à petit un milieu humain totalement nouveau. Les milieux ne sont pas des contenants passifs, mais des processus actifs. À l'âge électronique, un milieu totalement nouveau a été créé. Le "contenu" de ce milieu nouveau, c'est l'ancien milieu machiniste de l'âge industriel. Le nouveau milieu refaçonne l'ancien aussi radicalement que la télévision refaçonne le cinéma. Le contenu de la

⁶¹ R. Debray, *Cours de médiologie générale*, Gallimard, 1991.

⁶² D. Bougnoux, *Introduction aux sciences de la communication*, La Découverte, coll. Repères, 2002, p. 18.

⁶³ G. Bateson, *Vers une écologie de l'esprit*, Le Seuil, 1977.

⁶⁴ M. McLuhan, *Pour comprendre les médias*, Points, 1968, p. 21.

télévision, en effet, c'est le cinéma »⁶⁵. L'IA, en ce sens, se saisit de l'ancien monde (le monde pré-IA) et le refaçonne à sa guise. La célèbre formule de Lavoisier « rien ne se perd, rien ne se crée, tout se transforme » trouve ici à s'appliquer. Tout ce qui existait avant l'IA (les besoins, les questions, les réponses, les problèmes, les solutions, les dangers, les opportunités, les loisirs, le travail etc.) existe et existera encore dans le monde de l'IA, mais sous un jour différent, très différent, voire radicalement différent. Peut-être qu'après avoir longtemps relégué au second plan la question de la technique communicationnelle et informationnelle, il se pourrait que les anthropologues, sociologues et autres politologues en viennent à être hantés par elle. Chacun ressent bien à quel point les nouvelles machines informatiques et numériques reconstruisent les formes du savoir, le lien politique, le sentiment social et la culture en général.

Les technologies médiatiques ne piègent néanmoins pas des sujets pensants qui, à l'état de nature, seraient parfaitement indépendants, lucides, vigilants et éclairés. Au contraire, ces outils sont indispensables pour réfléchir et se protéger contre les risques d'aliénation propres à cet état de nature. La moindre information requiert, pour son extraction, son traitement et son acheminement, une technologie. Mais les IA dépassent très largement ces fonctions primaires de la technique informationnelle et vont jusqu'à indiquer à chacun ce qu'il doit penser, quelles doivent être ses idées et ses centres d'intérêt, cela en fonction de motivations intérieures ou extérieures à l'individu. L'homme, s'auto-inspectant, dira encore longtemps « *ego cogito*, je pense — c'est moi qui pense quand je pense, c'est moi qui parle quand je parle, c'est moi qui écrit quand j'écris ». Ses songes, ses paroles et ses écrits lui paraîtront forcément innés, libres, jaillissant du plus intime de son être. Cet homme ressemble pourtant de moins en moins au *Penseur* d'Auguste Rodin, ce monument d'idéalisme qui montre l'être humain nu de tout artifice, secouru par aucun livre ni aucun écran, seul à penser ses pensées. On n'apprécie guère, et il arrive même que l'on ne puisse pas, interroger les prothèses techniques qui formatent les discours. Peut-on étudier un instrument au moyen de cet instrument lui-même ? Cela ne revient-il pas à couper la branche sur laquelle on est assis ? Et il est vrai que le propre des médias, à commencer par le corps, est de rester en arrière-plan et de jouer leur rôle, si ce n'est de façon invisible, en tout cas en se faisant oublier. Mais l'influence de certaines IA sur le paysage médiatique, communicationnel et informationnel est telle qu'il devient urgent d'étudier le médium plus que le message — même si Marshall McLuhan aurait certainement observé combien ils tendent à se confondre toujours plus.

Par exemple, les mécanismes de filtrage des réseaux sociaux, alimentés en continu par les préférences exprimées par les internautes, finissent par les enfermer dans une certaine vision du monde qui s'auto-entretient et donc se renforce, sans possibilité d'être confronté à une quelconque contradiction. Alors qu'internet devait offrir un accès formidable à la diversité et à l'infinité du monde, les algorithmes des GAFAM et autres services de réseautage social, à la recherche de toujours plus de clics, enferment les individus sur eux-mêmes. L'algorithme d'un grand réseau social détermine ainsi quels contenus de quels « amis » doivent être affichés en priorité et aboutit à une hyper-personnalisation faisant que, en permanence, les textes, les images, les vidéos, les liens et les publicités qui s'affichent vont dans le sens des choix habituels des utilisateurs. Cette « autopropagande invisible »⁶⁶ favorise notamment les phénomènes de radicalisation. Le pluralisme des courants de pensée et d'opinion et, par suite, la liberté d'opinion sont mis en danger. Peut-être s'agit-il d'une véritable « dictature des algorithmes » tant leurs effets normatifs ne répondent guère

⁶⁵ *Ibid.*

⁶⁶ E. Pariser, *The Filter Bubble – What the Internet is Hiding from You*, Penguin Books, 2011. L'auteur donne l'exemple de deux personnes aux convictions politiques opposées qui recherchent au moyen d'un moteur de recherche des informations concernant une grande compagnie pétrolière. La personne de « droite » voit s'afficher des résultats se rapportant aux chiffres financiers et aux performances de l'entreprise, alors que celle de « gauche » obtient des résultats relatifs à la dernière marée noire et autres dégâts causés par cette compagnie. En affichant en priorité des éléments allant dans le sens de ce que les internautes aiment et pensent déjà, le moteur de recherche les emprisonne dans des « bulles cognitives ».

aux exigences démocratiques — loin du pluralisme externe de la presse (par la diversité des titres) et du pluralisme interne des chaînes de télévision et stations de radio (obligées par la loi de donner la parole à toutes les opinions sans défendre de ligne idéologique particulière). Les médias fonctionnant à base d'intelligence artificielle sont, eux, davantage des miroirs que des fenêtres.

C'est grâce à la pluralité des voix entendues, avec leur souhaitable cacophonie, et à la diversité des voies parcourues, avec leurs profitables chemins de traverse, que s'entretiennent l'esprit critique et la curiosité. Les chercheurs et les laboratoires savent à quel point la quête de la vérité a besoin de dispute et de contradiction et combien le plus grand danger est l'endormissement dans un sommeil dogmatique, dont il est parfois difficile de se réveiller pour aller s'abreuver à d'autres sources dont l'eau est *a priori* moins limpide. Emmanuel Kant, dans son opuscule *Qu'est-ce que les Lumières ?*, a montré qu'il n'est pas de plus sûr critère pour évaluer la vigueur d'une démocratie que celui de sa presse et de son pluralisme. Or les IA médiatiques ruinent la presse véritable, subrogée par le règne de la dépêche, de l'immédiateté, du ciblage et de l'absence de signature, et réduisent à la portion congrue le pluralisme. Face à ces nouveaux « nouveaux chiens de garde », qui « se dévoient dans une société de cour et d'argent, en se transformant en machines à propagande de la pensée de marché »⁶⁷, qui pourra stopper l'appauvrissement du débat public ?

La médiologie et, plus généralement, les sciences de l'information et de la communication ne peuvent donc qu'être fort préoccupées par les bouleversements subis par leurs objets en raison de l'intervention des IA médiatiques. Leurs spécialistes enseignent que « l'homme tient son humanité d'un certain régime symbolique, signifiant. Nous vivons moins parmi les choses que parmi une "forêt de symboles", comme dit Baudelaire, et ceux-ci nous rendent le monde familier en interposant entre lui et nous un ordre de signes, plus maniable et léger que celui des choses. L'empire des signes double ainsi notre monde naturel ; la sémiosphère contient la biosphère. Par tout un réseau de représentations codées et de signes qui sont autant de pare-chocs opposés à la dureté du monde, nous enveloppons, nous filtrons et du même coup nous maîtrisons le monde extérieur »⁶⁸. Certainement l'influence des IA sur cet empire des signes, sur cette forêt de symboles, sur ce réseau de représentations codées, sur la sémiosphère et, par suite, sur la biosphère est-elle immense, voire radicale dans certains cas. Qu'on pense ne serait-ce qu'aux algorithmes qui filtrent les contenus du web ou, dans un tout autre ordre d'idée, aux robots humanoïdes et aux IA qui visent à améliorer les relations homme-machine par la détection des émotions, le dialogue et la robotique affective et interactive.

L'homme descendant ainsi du signe plus que du singe, il a toujours souhaité avoir à sa disposition des médias (des intermédiaires cognitifs et relationnels) permettant de l'ancrer dans son port d'attache plutôt que de l'abandonner aux caprices des vents contraires, de le stabiliser plutôt que de l'égarer dans le vagabondage, de lui donner un agenda plutôt que de le laisser face à l'inconnu, en bref de le fixer dans son espace et son temps propres, rassurants parce que familiers. Mais, de tout temps, l'état des technologies ne permettait pas une telle « pertinence » des informations et, par suite, la balkanisation des médias. Avec les IA, au contraire, la personnalisation des contenus atteint son paroxysme. Et « ma culture » devient « ma clôture ».

Dans le royaume des *nudges*⁶⁹, lorsque l'homme ne prend plus de décisions éclairées — même s'il prend peut-être les bonnes décisions —, qu'il ne comprend plus pourquoi il agit dans un sens plutôt que dans un autre, qu'il n'a plus entièrement conscience de ses actes, de leurs causes et de leurs finalités, il se robotise. Tel pourrait être le grand paradoxe de l'humanisation des robots au moyen de l'intelligence artificielle : robotiser les humains. Pour éviter cela, il faut comprendre et

⁶⁷ S. Halimi, *Les nouveaux chiens de garde*, Liber-Raisons d'agir, 1997, p. 159.

⁶⁸ D. Bougnoux, *Introduction aux sciences de la communication*, La Découverte, coll. Repères, 2002, p. 28.

⁶⁹ En sciences du comportement, la théorie des *nudges* (ou théorie du paternalisme libéral) explique que des suggestions indirectes peuvent, sans forcer, influencer les motivations et les prises de décisions individuelles ou collectives souvent plus efficacement que des instructions directes, des normes, des lois.

expliquer au plus grand nombre les algorithmes et les big data, donc les vulgariser, les démocratiser. Rarement le besoin d'une science grand public s'est fait autant ressentir car rarement des progrès scientifiques ont impacté aussi intimement le quotidien de chacun⁷⁰.

Les intelligences artificielles sont forcément politiques et même idéologiques, jamais neutres. On fabrique ces extraordinaires outils afin qu'ils produisent des significations différentes, des cadres originaux, des routes nouvelles, dessinant le monde de demain, un monde tout autre de celui d'hier. Les IA façonnent un environnement technologico-politique particulier dans lequel chacun est amené à vivre, à penser et à juger autrement. La société pourrait n'être bientôt plus qu'une société de calculs, réduite en nombres, en fonctions mathématiques, en statistiques et en indicateurs. À partir du passé, les calculateurs définissent et organisent la réalité d'aujourd'hui et de demain. Ils produisent des conventions et des systèmes d'équivalence qui choisissent certains objets et certaines valeurs au détriment des autres, redessinant progressivement et insidieusement les cadres cognitifs et culturels des sociétés contemporaines. Les algorithmes, quand ils ne sont pas réglés afin de servir l'« autopropagande invisible » et enfermer chacun sur lui-même, cherchent à enfermer chacun dans certains environnements bien choisis, cadrer les conduites, formater les esprits, décider des désirs, standardiser les besoins, donc favoriser le suivisme, le panurgisme, et éliminer au maximum les envies et les modes de vie alternatifs — ceux-ci étant trop peu monétisables. Les utilisateurs se retrouvent donc placés dans des silos de comportement et de consommation tracés par les algorithmes. Quand la normalité devient normativité.

Tous ces calculs ne sont possibles que parce que les individus modifient leurs comportements afin de les rendre calculables. Le problème est qu'ils le font souvent de façon inconsciente et involontaire. Les technologies ne fonctionnent que parce qu'elles opèrent dans un milieu associé qui les rend efficaces et pertinentes. Pour éviter de se laisser conquérir par l'intelligence artificielle, afin de ne pas permettre à la dictature des algorithmes de s'imposer, et pour, au contraire, s'allier à ces technologies, il faut comprendre quelle est la situation, à quel point les robots peuvent suggérer et décider de beaucoup de choses. Ce n'est qu'ainsi que chacun pourra préserver sa souveraineté individuelle, veiller sur sa liberté d'autodétermination, demeurer maître de soi-même dans un contexte où les IA et certaines puissances (privées ou publiques) dont elles sont les instruments cherchent à limiter le libre-arbitre et la diversité.

L'efficacité et le confort des intelligences artificielles, omniprésentes dans les décisions et les conduites quotidiennes, tendent à endormir le sens des responsabilités, l'autonomie et l'individualité⁷¹. Or ne faudrait-il pas conserver, pour reprendre une expression de John Stuart Mill, un « droit à l'errance », une capacité à dévier, à délibérer intérieurement, à innover, à s'autogouverner, à choisir et à se choisir, en somme à être libre⁷² ? Revient l'ancestral débat libre arbitre contre déterminisme. En amenant les hommes à de moins en moins réfléchir, penser, garder en mémoire et prendre des décisions, l'intelligence artificielle risque de nuire à l'intelligence humaine. Cela conduit, par exemple, Éric Sadin à s'inquiéter du pouvoir injonctif des IA, lequel serait profondément antihumaniste⁷³. À des fins de normalisation et d'optimisation, on formate et

⁷⁰ En ce sens, Cédric Villani a souhaité que son rapport de 2018 s'adresse et soit accessible à tout le monde, soulignant combien le sujet concerne et touche tout le monde.

⁷¹ G. Koenig, *La fin de l'individu – Voyage d'un philosophe au pays de l'intelligence artificielle*, Éditions de L'Observatoire, 2019.

⁷² Par exemple, à l'ère de la voiture autonome, obligera-t-on la population à rouler en voiture automatique-connectée ou acceptera-t-on encore les conducteurs traditionnels, bien que cela ne soit pas profitable à la communauté, à l'intérêt général, car ils provoquent des accidents et une gestion plus difficile du trafic, créant des situations non optimales pour le reste du groupe. Le premier choix est utilitariste, en faveur du collectif, le second est individualiste, protecteur des libertés individuelles en préservant la possibilité d'exprimer une personnalité singulière y compris sur la route.

⁷³ É. Sadin, *L'intelligence artificielle ou l'enjeu du siècle – Anatomie d'un antihumanisme radical*, L'Échappée, 2018. Le philosophe distingue quatre niveaux parmi les injonctions technologiques, selon leur plus ou moins grande capacité à décider à la place de l'humain : le niveau « incitatif » (avec l'exemple du GPS, qui laisse une certaine marge de manœuvre aux conducteurs), le niveau « impératif » (lorsqu'il s'avère plus difficile de ne pas suivre les indications des

limite la faculté de jugement et d'action, on y substitue des protocoles informatiques afin d'infléchir dans un sens prédéterminé les actes, de réduire la liberté, de décider à la place de chaque individu. Ainsi s'agirait-il d'éviter les erreurs, la déviance, cela dans la veine d'une idéologie selon laquelle, Dieu n'ayant pas terminé la création, les hommes devraient être corrigés de leurs défauts, remis dans le droit chemin. Les nouvelles technologies seraient très utiles pour cela en permettant de parer à nombre de travers humains fort coûteux, notamment socialement.

Une telle vision utilitariste, collectiviste et même hygiéniste n'est évidemment assumée par personne. Elle paraît pourtant bel et bien empreindre pour une part l'environnement idéologique dans lequel les IA se développent. Celles-ci servent un mouvement global d'hyper-rationalisation des sociétés, dans lesquelles on veut à tout instant prendre les meilleures décisions afin de tout optimiser et tout maximiser. Au-delà, et pour la première fois de l'histoire, les vérités — avec leur valeur idéologique, distinctes des réalités matérielles — sont définies par des machines, par des technologies jouissant d'une immense puissance performative.

Il n'est pourtant pas rare, et même plutôt fréquent, que le merveilleux, l'agréable ou le profitable provienne de l'incertain, de l'aléatoire, du naturel, de l'instinctif ou de l'impulsif. L'ordre anarchique et l'ordre spontané peuvent paraître plus beaux, plus attirants, plus séduisants, plus agréables que l'ordre parfait entièrement programmé, mathématique, géométrique et logique — il suffit d'observer quels quartiers des grandes villes attirent le plus de touristes : il s'agit généralement des centres désordonnés et restés sous une influence moyenâgeuse, qui sont très peu pratiques mais qui présentent un charme certain. Le « droit à l'errance », foncièrement humain, est aussi un « droit à l'erreur », un « droit à la déviance », un « droit à la divagation » et un « droit à la sérendipité », lesquels sont des conditions du progrès. Le risque qui accompagne la standardisation des comportements sous l'influence des IA est celui de l'immobilisme⁷⁴ — par exemple, la *smart city* serait une ville morte, sans caractère et sans humanité. L'optimisation peut être source d'inertie. Cette inertie, sociale et politique, pourrait être le résultat de la diminution du pluralisme des hommes, de la réduction de la diversité de l'humanité. La pluralité et la liberté, sous toutes leurs formes, ne sont-elles pas les conditions essentielles de l'autonomie de jugement et, par suite, de la démocratie, de la politique, du progrès et de bien d'autres choses ? La prophétie technocratique saint-simonienne du « remplacement du gouvernement des hommes par l'administration des choses » pourrait-elle se réaliser dès lors que les hommes auraient évolué en choses, que le matériau pétri par le politique serait devenu inerte, que les gouvernants auraient face à eux des sujets sans inspiration et sans aspirations, pauvres en idées et en volonté, dépourvus d'esprit critique et incapables de proposer des projets alternatifs ou de soutenir des discours discordants ? Il importe de conserver la possibilité, à tout instant, d'inventer et de réinventer le monde et son monde autrement, à la fois à titre individuel et collectivement.

Tout cela interpelle l'historien israélien Yuval Harari, qui se demande si les applications industrielles de l'IA, qui en viennent à connaître les individus mieux que les individus eux-mêmes se connaissent, ainsi que les *nudges* de plus en plus omniprésents, ne conduisent pas à déléguer la capacité de choisir et de décider à des machines, donc à s'y soumettre dans des pans entiers de la vie

IA, à l'image de ces *chatbots* qui font passer des entretiens d'embauche et indiquent à l'humain, qui est encore dans la boucle, qui il doit recruter, ou des exemples des algorithmes des banques, de la justice prédictive ou de la police prédictive), le niveau « prescriptif » (notamment dans le cas de la médecine, où il est quasiment impossible de ne pas suivre les diagnostics de l'IA tant cela semble risqué) et le niveau « coercitif », le plus impactant, quand l'homme est totalement obligé par la machine, sans plus aucune liberté d'action propre (par exemple dans les domaines de l'économie et du travail où des mégadonnées sont analysées en temps réel et permettent à des algorithmes de passer des commandes à des manufacturiers et à un personnel qui ne peut alors que s'exécuter face à des contremaîtres insensibles et informels).

⁷⁴ Il a été montré que les chauffeurs qui utilisent le GPS toute la journée se robotisent et sont atteints d'une atrophie de certaines zones cérébrales à force de ne plus faire de petits choix et d'agir de manière automatique, par réflexes, tout au long de la journée.

individuelle, sociale et professionnelle⁷⁵. La mathématicienne et militante américaine Cathy O'Neil déplore en ce sens que des « armes de destruction mathématiques » (« *Weapons of Math Destruction* ») décident des affectations des étudiants dans les universités, accordent les crédits, fixent les habitudes de consommation, attribuent les emplois et, surtout, déterminent les votes lors des élections⁷⁶. Il serait donc impératif de prendre conscience et de se protéger de l'emprise et des dérives des algorithmes, rendues possibles par la connexion permanente des individus et par la puissance de calcul toujours plus importante des ordinateurs. Et ce sont les universitaires spécialistes de ces questions qui doivent jouer le rôle clé dans cette prise de conscience et cette protection.

Cela semble d'autant plus indispensable que les nouvelles technologies permettent beaucoup plus facilement de créer et maintenir des régimes totalitaires, bien que la situation ne soit pas tout à fait celle décrite par George Orwell dans *1984* : il n'y a pas, ici, de super-gouvernement qui contrôlerait les habitants de l'extérieur, la liberté d'expression ayant disparu et toutes les pensées étant minutieusement surveillées, tandis que des affiches placardées dans les rues indiqueraient à tous que « Big Brother vous regarde » (« *Big Brother is watching you* »). Il s'agit davantage d'un autocontrôle pénétrant jusque dans la plus stricte intimité, chacun fournissant les moyens de son propre contrôle⁷⁷. Le résultat est cependant à peu près le même : l'avènement d'une société de la surveillance et de réduction des libertés. La performativité et les prophéties auto-réalisatrices des IA sont telles que, de fait, la technologie choisit pour l'humain qui s'en remet assez aveuglément à elle. Ce « coup de data permanent » enferme les individus dans des silos de comportement, de consommation, mais aussi de pensée. Cela porte atteinte à la liberté d'opinion et au pluralisme des courants d'idées. Se pose dès lors la question de la capacité de la société numérique à être une société démocratique et ouverte plutôt qu'une société tyrannique et fermée. Mais chacun peut agir afin de changer les choses, à son échelle personnelle puis à une échelle collective. Et les chercheurs ont les moyens de devenir les moteurs de cette rétroaction contre le *diktat* algorithmique, non en manifestant une opposition de principe et catégorique mais en livrant au public les clés pour bien comprendre les mutations en cours, notamment en matière d'orientation des conduites individuelles et collectives.

Il faudrait n'abandonner aux robots que ce qui déshumanise et qui entrave pour retrouver le temps d'apprendre et de s'apprendre, de comprendre et de se comprendre, de connaître et de se connaître ; n'utiliser ces nouveaux outils que lorsqu'ils créent de l'indépendance et non de la dépendance ; et conserver ce qui libère, ce qui épanouit. L'intelligence artificielle, neutre en soi, pourrait aussi servir une grande transformation humaniste, comme l'écriture ou l'imprimerie, avant elle, ont permis de grands progrès de l'instruction et de la connaissance. Ne confondant pas le plus et le mieux, on passerait d'une civilisation du plus à une civilisation du mieux, réflexive, patiente et critique.

Au-delà de la compétition scientifique mondiale, l'intelligence artificielle est devenue l'affaire de tous. Dans ces conditions, les chercheurs doivent plus que jamais se mettre au service de la population, diffuser les connaissances, vulgariser ce qui présente la plus grande complexité. Leur importance est d'autant plus grande que, s'agissant d'un sujet déterminant pour l'avenir de

⁷⁵ Y. Harari, *Homo Deus – Une brève histoire de l'avenir*, Albin Michel, 2017.

⁷⁶ C. O'Neil, *Weapons of Math Destruction – How Big Data Increases Inequality and Threatens Democracy*, Crown, 2016 ; *Algorithmes – La bombe à retardement*, Les Arènes, 2018. Il faut toutefois noter que Cathy O'Neil cite essentiellement des exemples et anecdotes américains et que nombre des situations décrites sont supposées ne pas pouvoir se produire en Europe, où le droit est quantitativement et qualitativement plus protecteur des individus, notamment de leur vie privée et de leurs données personnelles.

⁷⁷ Lors de son intervention au cours d'un colloque au Sénat sur l'intelligence artificielle le 19 janvier 2017, Laurent Alexandre expliquait ainsi que « nous sommes tous les “idiots utiles” de l'IA en mettant chaque jour des milliers d'informations au service des plateformes du numérique » (référence à Lénine qui aurait utilisé cette expression afin de désigner les intellectuels de gauche occidentaux défendant avec enthousiasme et naïveté le régime soviétique).

l'humanité, il est nécessaire de pouvoir en débattre en toute indépendance, sans défendre des intérêts technologiques, politiques ou économiques. La question nietzschéenne « qui parle ? », à travers l'individu qui s'exprime, est incontournable et essentielle. Le grand public peut accéder à des informations objectives et à des avis éclairés lorsque le discours n'est pas orienté par de quelconques intérêts, ce qui peut être le cas de celui des chercheurs — mais ne l'est pas *ipso facto*, d'autant plus que des convictions, des idéologies ou même ses propres écrits antérieurs sont d'autres formes d'intérêts. Au-delà, les scientifiques doivent également s'adresser aux gouvernants (publics et privés), afin de les aider à faire les bons choix, des choix structurants au moment où l'on prend des virages décisifs.

La grande aventure de l'intelligence artificielle est essentiellement un enjeu d'intelligence humaine : celle des chercheurs, celle des développeurs et des ingénieurs, celle des entrepreneurs et des investisseurs, celle des gouvernants, mais aussi celle de chaque personne qui, tous les jours, toutes les heures, à tout instant, agit à l'aune de propositions automatiques. Chacun doit pouvoir être acteur, et non seulement spectateur ou objet, de l'épopée de l'intelligence artificielle ; avec en toile de fond l'idée-guide selon laquelle le progrès technologique n'a de sens que s'il accompagne et stimule un progrès social et un progrès humain, que s'il bénéficie aux hommes et à leur capacité à s'organiser en société pacifique et prospère⁷⁸.

Peut-être les français et les européens sont-ils ceux qui, dans le monde, attachent le plus d'importance à ces objectifs humains et sociaux. En témoigne le slogan anglophone attaché par le gouvernement français à sa stratégie en intelligence artificielle : « *AI for Humanity* ». C'est ainsi que, sur le « vieux continent », on a tendance à se focaliser sur la « robéthique », à investir dans des comités et des rapports éthiques, à se spécialiser dans les précautions morales et juridiques plutôt que dans les avancées technologiques et scientifiques, au risque de subir une « fuite des cerveaux » et une perte de puissance économique, technologique et politique⁷⁹. Doit-on le regretter ? Si le monde se divise entre ceux qui créent des IA et ceux qui créent des comités d'éthique sur les IA, de quel côté vaut-il mieux se trouver ? Ces comités ont sans doute toute leur raison d'être s'il faut se méfier à la fois des innovations qui ne sont dictées que par l'appât du gain et de la politisation hystérique de la science, le tout dans un contexte mêlant, en fonction du point de vue adopté, crise de la rationalité et rationalisation extrême. Toute recherche qui avance sans se préoccuper de ses conséquences, directes et indirectes, peut être dangereuse et ne pas aller de pair avec le progrès.

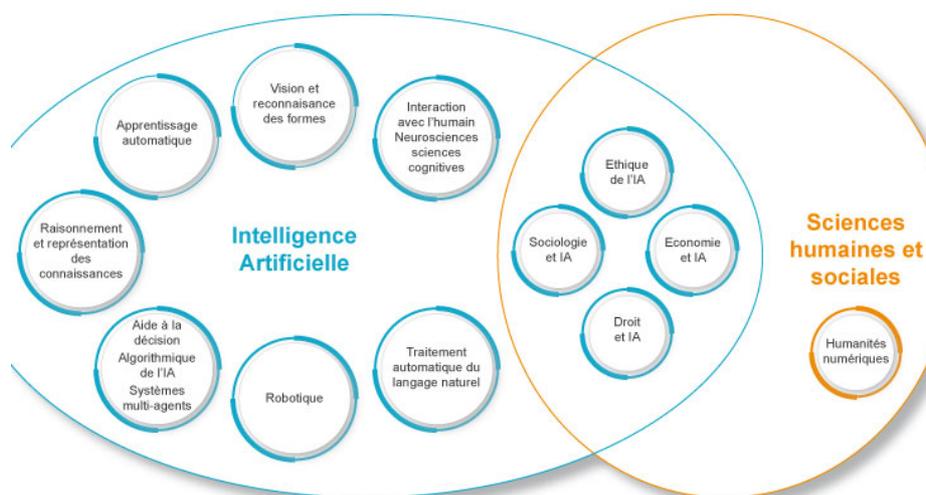
L'Europe pourrait-elle être en retard en matière de technologie de l'IA et d'économie de l'IA, mais en avance en matière de pensée de l'IA et d'éthique de l'IA ? Sous bien des angles, celles-ci ne sont pas moins importantes que celles-là. L'IA a besoin de chercheurs travaillant sur les questions éthiques, réfléchissant à son rapport aux droits et libertés fondamentaux, à la Constitution et à la démocratie. Le siècle des intelligences nécessite un droit des intelligences. En témoignent ces groupes de plus en plus nombreux de juristes et d'informaticiens qui attaquent les projets de textes de droit liberticides du point de vue des « libertés numériques », à une époque où chacun se promène en permanence avec dans la poche un micro-ordinateur-espion muni de deux caméras, d'un micro et d'une ribambelle de programmes ultra sophistiqués de surveillance et de traçage. Le rapport entre le couple droit-politique et l'intelligence artificielle est d'autant plus important — et intéressant — que, si le droit et la politique saisissent l'IA (statut juridique, propriété intellectuelle, responsabilités etc.), il est tout aussi vrai que l'IA saisit le droit et la politique et, parfois même, fait le droit et la politique (*legaltechs*, *civictechs*, performativité des algorithmes etc.).

⁷⁸ En ce sens, il y a plus de puissance informatique dans un smartphone d'aujourd'hui que dans tout le projet Apollo qui a permis d'envoyer des hommes sur la Lune. Or personne, avec son smartphone, n'est capable de se rendre sur la Lune, car il manque l'intelligence collective, la volonté, la motivation, la détermination, autant de facteurs humains nécessaires au-delà de la technologie qui, en soi, ne peut rien.

⁷⁹ En ce sens, dans son rapport de mars 2018, Cédric Villani propose de créer « une instance pour émettre des avis, donner des jugements en toute indépendance, qui puisse être saisie par le gouvernement comme par les citoyens, et qui nous dise ce qui est acceptable ou non acceptable ».

Mais peut-être, au-delà du droit, des normes, de l'éthique et de la régulation, des mobilisations individuelles et collectives, pour l'instant quasi-inexistantes, sont-elles nécessaires afin de prévenir certaines dérives de l'IA ou simplement mieux la comprendre, mieux la saisir de différentes manières. Éric Sadin invite en ce sens ses lecteurs à « mettre en crise » les nouveaux modèles en train de s'affirmer en refusant les actes d'achats de produits générateurs de mégadonnées (balance connectée, tétine connectée, réfrigérateur connecté, bikini connecté etc.)⁸⁰. Et Régis Debray a souligné combien, tandis que la banalisation de l'automobile a tendance à réduire drastiquement la marche à pied, les individus qui ne marchent plus se mettent à courir, comportement spontané, sauvage, naturel, quand la marche est civilisée. Il a baptisé ce retour à l'archaïsme « effet jogging »⁸¹. Se pourrait-il que les développements de l'IA s'accompagnent de quelques « effets jogging » — à l'image de ces internautes qui consacrent leur temps à référencer des pages et des liens du web dans un univers numérique où les algorithmes des moteurs de recherche rendent déjà tous les services possibles et imaginables ?

Alors que les spécialistes de la technologie, dans la Silicon Valley comme ailleurs, peuvent se laisser absorber par les seules questions des performances et de l'efficacité des outils qu'ils élaborent et qu'ils cherchent à améliorer, d'autres doivent interroger les conséquences sociales, économiques, politiques, éducatives, éthiques, scientifiques et parascientifiques (conséquences des progrès d'une science sur les autres sciences). L'IA, en soi, est bien sûr une affaire de mathématiciens, d'informaticiens, de neuroscientifiques et de roboticiens. L'économie, le droit, la sociologie ou l'anthropologie sont des sciences périphériques à l'intelligence artificielle. Elles n'en sont pas moins indispensables, cela afin de la comprendre mais aussi afin de la guider et même la diriger.



C2RP, septembre 2017

Ce besoin de scientifiques des IA d'origines diverses et uniquement motivés par la meilleure compréhension des phénomènes socio- technologiques à l'œuvre ou à venir deviendra toujours plus important avec le développement de l'homme augmenté et du transhumanisme. Pour l'heure, l'homme n'est que légèrement augmenté, avec un téléphone-ordinateur portable, une connexion

⁸⁰ É. Sadin, *La silicolonisation du monde – L'irrésistible expansion du libéralisme numérique*, L'Échappée, coll. Pour en finir avec, 2016.

⁸¹ R. Debray, *Introduction à la médiologie*, Puf, coll. Premier cycle, 2000. Le philosophe observe ainsi qu' « au début du siècle, certains visionnaires avaient pronostiqué que l'usage immodéré de l'automobile par les citoyens provoquerait bientôt l'atrophie de leurs membres inférieurs, le bipède motorisé se désaccoutumant de la marche. Qu'a-t-on vu depuis ? Ceci : depuis que les citoyens ne marchent plus, ils courent. Fanatiquement. Dans les parcs ou, à défaut, en salle, sur tapis roulant ».

internet permanente, une montre connectée, un thermostat intelligent etc. Mais, s'il en venait à utiliser des lentilles de contact connectées, à s'implanter des puces dans le corps et dans le cerveau, démultipliant ses capacités cérébrales, ou si des manipulations génétiques permettaient de ne plus donner naissance qu'à des enfants au quotient intellectuel très élevé, des questions éthiques autrement complexes se poseraient. La fusion de l'intelligence humaine et de l'intelligence artificielle⁸², que certains appellent de leurs vœux, pourrait s'analyser tel un progrès ou telle une régression, si ce n'est tel le début de la fin de l'humanité. Et « IA » seraient les initiales du nouveau dieu, du nouveau guide suprême de l'humanité⁸³.

« Homme augmenté », « intelligence augmentée », « *brain editing* », « *neurohackers* », « neuroprotecteurs », « neuro-intimité », paramétreurs d'implants intra-cérébraux, entrée du cerveau dans l'internet des objets, eugénisme technologique⁸⁴... En pompiers pyromanes, certains sont tentés de fusionner avec le feu. Quelques garde-fous semblent alors ô combien utiles. Et les chercheurs semblent être les mieux en mesure de produire et défendre ces garde-fous, parce qu'ils sont (normalement) indépendants, parce qu'ils sont les mieux informés des enjeux et des potentialités positives comme négatives de la technologie qu'ils étudient, et parce qu'ils sont les mieux en mesure de prendre du recul — ce qui est indispensable pour prendre de l'élan et aller de l'avant.

Il n'y a cependant pas de chemin allant directement de la connaissance de ce qui est à la connaissance de ce qui doit être. Beaucoup de choix de société ne sauraient être dictés ni par des algorithmes, aussi perfectionnés et efficaces soient-ils, ni par des scientifiques, aussi objectifs et insubordonnés soient-ils. Bien sûr, toute morale est contingente et relative, variant dans le temps et dans l'espace, ainsi qu'entre les individus au sein d'une même population. L'appréhension morale de la technique est donc difficile. La question de l'IA et de la morale est même sans doute insoluble. Il y a forcément des conflits de morales qu'on ne peut pas trancher, si ce n'est arbitrairement, sans pouvoir soutenir son choix par une argumentation assurément supérieure. Tel est le cas, par exemple, concernant les accidents de voitures autonomes : qu'est-ce qui est le plus moral entre écraser trois seniors et écraser deux enfants ?

Ces difficultés ne sauraient pour autant justifier de se retrancher derrière les objets les plus objectifs et les problématiques les moins polémiques, ne sauraient inciter à se cantonner à la simple analyse technologique de l'IA. Formellement et matériellement, il faudra trouver les moyens de réguler et non laisser faire les robots, les hommes augmentés et le nouvel « *homo deus* ». On pourrait, par exemple, craindre la prolifération d'une forme de racisme cognitif : une discrimination

⁸² Le 28 avril 2017, Elon Musk a annoncé le lancement de Neuralink, une startup neurotechnologique qui développe des implants cérébraux d'interfaces neuronales directes. L'objectif est d'hybrider les cerveaux, notamment afin d'éviter que l'homme ne se laisse dépasser par la machine et devienne son domestique. Il s'agit donc d'aider les 83 milliards de neurones de l'homme grâce à des circuits miniatures permettant d'augmenter son quotient intellectuel et de le rendre plus compétitif face à l'IA. De son côté, Facebook a annoncé travailler sur une technologie de télépathie devant permettre de transférer un message d'homme à ordinateur ou d'homme à homme par la pensée, à la vitesse de 100 mots par minute.

⁸³ Aux États-Unis, Anthony Levandowski, ingénieur et père de la voiture autonome de Google, pensant que l'intelligence artificielle dépassera largement et en tout l'intelligence humaine, propose de lui vouer un culte. Il a ainsi fondé une organisation religieuse baptisée « *Way of the Future* » (« La voie de l'avenir ») qui fait la promotion d'une « divinité » basée sur une intelligence artificielle et qui est censée permettre le progrès de la société. Cela accreditte l'idée que le transhumanisme serait bien un discours religieux plutôt que scientifique, profitant de la tentation perpétuelle de l'humanité à vouloir se sauver par ses propres forces et de l'attrait pour le gnosticisme, qui consiste à voir dans le corps une entrave et à rêver de s'en débarrasser. L'âme serait détachable du corps et pourrait être mise sur un microprocesseur.

⁸⁴ En intervenant sur l'ADN, avec des manipulations génétiques, en augmentant technologiquement les hommes, on pourrait créer des hommes sur mesure, à la carte — par exemple de très bons musiciens ou sportifs, mais aussi des militaires sans peur, sans compassion, sans regrets et sans douleur. Peut-être est-ce plus effrayant que la bombe nucléaire. Que devra faire l'Europe si des pays, ailleurs dans le monde, décident d'exploiter ces possibilités technologiques et scientifiques ?

et un rejet des plus intelligents et des augmentés envers les moins intelligents et les non augmentés. Et puis, avec la pantoprédicibilité et la dictée sociale des algorithmes, il faudrait prévenir la tentation des générations futures de tomber dans la datacratie, l'algocratie, le dataisme et la datadépendance, c'est-à-dire l'enfermement dans un monde de chiffres, de statistiques et d'algorithmes déshumanisants et dénaturants. Peut-être mieux vaut-il conserver une grande part de hasard, lequel définit depuis toujours la vie et la marche du monde.

Les bioconservateurs pourraient-ils triompher et interdire ou supprimer les technologies ? Ou bien, tout à l'inverse, l'anarchie technologique pourrait-elle l'emporter, avec un développement tous azimuts uniquement motivé par la quête de nouveaux marchés, de nouveaux profits et de nouveaux monopoles, peu important que cela nuise à l'humanité et à l'intérêt général ? Peut-être la meilleure voie à suivre est-elle, comme souvent, celle de l'équilibre, ici en encadrant raisonnablement le pouvoir démiurgique de l'homme et ses relations avec l'IA. Au Canada, par exemple, on revendique des politiques en intelligence artificielle faisant la part belle à la liberté d'entreprise, comme aux États-Unis, tout en insistant sur les enjeux éthiques, comme en Europe. Le mot d'ordre est là-bas « *non predatory* », les progrès technologiques n'étant autorisés qu'à condition de servir l'intérêt public.

Pour aborder sans détours et sans biais idéologiques les dilemmes d'aujourd'hui et, sous un angle prospectiviste, agir sur le cours du futur en train de se faire, il semble nécessaire que la parole soit donnée aux savants — ou qu'ils la prennent. En ce sens, on ne peut que se réjouir du fait que, comme l'indique le rapport de Cédric Villani de mars 2018, la France est le quatrième pays au monde à produire le plus d'articles scientifiques sur l'intelligence artificielle, après la Chine, les États-Unis et le Royaume-Uni. Ces articles sont écrits par des mathématiciens, mais aussi par des spécialistes des sciences de l'information et de la communication ou des sciences humaines et sociales⁸⁵. L'explosion de l'intelligence artificielle s'accompagne donc — et c'est heureux — d'une explosion des sciences de l'intelligence artificielle.

Ce boom scientifique s'est produit à partir de 2012-2013, au moment du printemps de l'apprentissage profond et des réseaux de neurones :

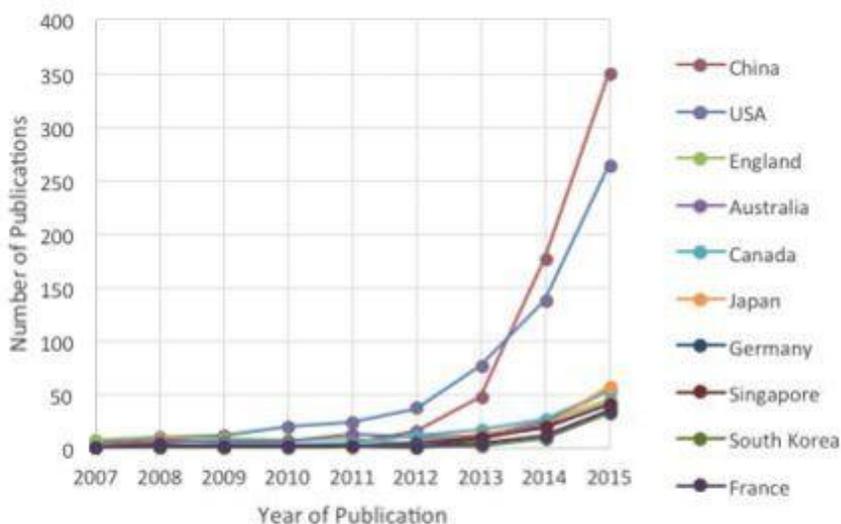


Figure 1: Journal articles mentioning "deep learning" or "deep neural network"

⁸⁵ Mais ces travaux universitaires ne sont pas suivis de réalisations concrètes ni de développements économiques et le rapport de Cédric Villani constate que « la France dispose d'un savoir théorique indéniable avec des cerveaux recherchés à l'étranger, mais qui ne se traduit pas encore par l'émergence de grandes entreprises leaders ». Les États-Unis, la Chine, le Royaume-Uni, le Canada et Israël sont les pays les plus avancés en termes de déploiement de l'intelligence artificielle.

L'avenir est investi de toutes sortes de peurs, manipulé par le personnel politique à des fins électoralistes, tandis que le politiquement correct interdit d'affronter (et parfois même de voir) certains problèmes et certains enjeux, en premier lieu en matière d'intelligence — quoique l'IA soit susceptible d'amener bientôt à remplacer le tabou de l'intelligence par une obsession de l'intelligence⁸⁶. Le monde politique n'est de toute façon pas bien armé car il n'a jamais eu à traiter des progrès exponentiels tels que ceux de l'IA⁸⁷ et parce qu'il fonctionne trop à base de démagogie et de court-termisme. Il en résulte un problème de désynchronisation et de rythme : l'IA et les technologies NBIC⁸⁸, technologies de l'exponentiel, progressent très vite, quand les institutions, notamment politiques et éducatives, ne changent guère, apparaissant parfois même sclérosées. Dans ces conditions, il appartient peut-être aux chercheurs — des universités et au-delà — de prendre le relai et de se saisir pleinement de ces sujets, dans toutes leurs dimensions.

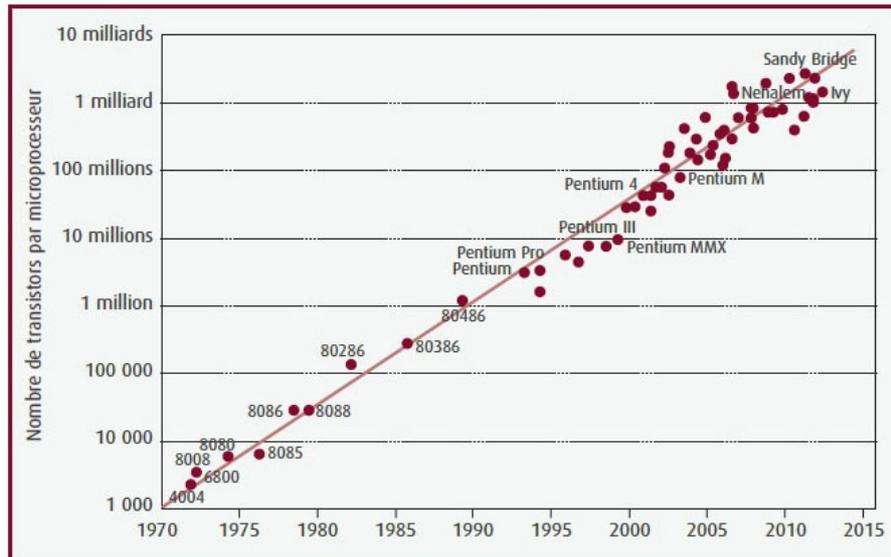
Leurs études doivent être substantiellement et sérieusement prospectives. En France, les travaux en prospective sont peu considérés et, par suite, peu développés. Rares sont d'ailleurs les prospectivistes-scientifiques, objectifs, honnêtes et non-fantaisistes, et la question des lendemains technologiques se retrouve presque en situation de jachère intellectuelle. Il faudrait pourtant pouvoir agir en fonction de ce que l'on veut et décider de ce que l'on veut en fonction de ce que l'on sait et non en fonction d'utopies ou dystopies, de mythes, de fantasmes ou de peurs infondées. En matière d'intelligence artificielle, l'écart entre ce qu'en montre Hollywood — et donc l'image que le grand public s'en fait — et ce qu'est sa réalité est aujourd'hui abyssal. Plus les futurs possibles seront précisément identifiés et réalistes, plus l'on pourra opérer un choix éclairé parmi eux et agir afin que le futur choisi advienne, donc prendre aujourd'hui les meilleures décisions et mesures politiques et juridiques. Cependant, le vrai pouvoir tend à échapper aux pouvoirs publics et à migrer vers les technologies, les infrastructures et donc ceux qui les possèdent. Aussi faudrait-il pouvoir intervenir à leur niveau plutôt qu'à celui des pouvoirs « officiels ».

La prospective de l'intelligence artificielle est toutefois rendue difficile par le fait que les progrès exponentiels des technologies NBIC les rendent largement imprévisibles, y compris dans leurs conséquences sociales et économiques. Par exemple, le progrès soudain de l'IA et des réseaux de neurones à partir de 2012 n'avait été anticipé par personne. Les derniers processeurs comportent sept milliards de transistors et effectuent un milliard d'opérations par seconde contre une opération toutes les 30 secondes pour la calculatrice mécanique de Blaise Pascal. La « loi de Moore » pourra-t-elle se prolonger encore longtemps ?

⁸⁶ Il n'est en effet pas politiquement correct de dire, et donc personne ne le dit, que ce sont les individus les plus intelligents qui prennent le pouvoir politique, économique, scientifique, technologique, financier etc. L'intelligence est la principale source de pouvoir. Plus l'IA se développe, plus il devient impératif d'accepter de parler de l'intelligence, pour défendre, promouvoir et développer les spécificités humaines-biologiques-naturelles.

⁸⁷ Il a fallu 550 millions d'années pour passer du premier neurone au premier transistor, inventé en 1947, puis seulement 70 ans pour en arriver aux technologies d'apprentissage profond. La puissance informatique a été multipliée par 95 millions de milliards en 70 ans.

⁸⁸ Les technologies NBIC sont les nanotechnologies, les biotechnologies, l'informatique et les sciences cognitives.



Un autre frein à une meilleure appréhension du futur, qui est en même temps un défi à relever pour les scientifiques souhaitant s'adresser au grand public, est qu'aujourd'hui les mêmes canaux de communication diffusent les connaissances, l'information, la désinformation, les commentaires, les croyances et les opinions. D'ailleurs, y compris les chercheurs s'expriment sur les réseaux sociaux. Ces divers contenus, qui ont des qualités objectives et des statuts différents mais qui se confondent dans des contenants identiques, sont souvent placés au même niveau et assimilés les uns aux autres par le grand public. Ainsi des amalgames, des brouillages, des incompréhensions et de la malinformation se produisent-ils. Par conséquent, l'enjeu pédagogique n'a de cesse de se renforcer. Il est dommageable que les divers statuts des contenus se contaminent ou même fusionnent. Pour éviter cela, il faudrait mieux connaître nos connaissances afin de pouvoir les expliquer en cas de discussion avec quelqu'un qui oppose des discours non scientifiques mais jouant sur les sentiments et les peurs des populations pour convaincre, à base de discours spectaculaires et démagogiques.

Toujours est-il que Sergey Brin, co-fondateur de Google, en vient à promettre la création prochaine de « machines qui raisonnent, pensent et font les choses mieux que les êtres humains ». Si le réalisme de pareille prophétie n'est guère assuré, il est en tout cas certain que l'IA, ensemble de technologies fondamentalement transversal et dont les limites du potentiel restent inconnues, impacte et impactera l'ensemble des disciplines. Elle remodèle et remodèlera la politique, l'économie, l'éducation, l'industrie, la santé, la police, la culture, le journalisme, la philosophie, le droit etc. Elle oblige et obligera à réinterroger ce qu'est et comment fonctionne une société ; et à réinterroger ce qu'est et comment fonctionne un homme.

En somme, l'intelligence artificielle est le grand sujet d'avenir et il faudra attendre longtemps pour que les recherches en la matière s'essoufflent. La problématique de la cohabitation entre intelligence biologique et intelligence technologique, entre intelligence naturelle et intelligence artificielle, est installée pour des décennies voire des siècles. Le plus passionnant est que tout ou presque reste à faire, à dire et à écrire. Le futur est incertain, vertigineux, extraordinaire. Pareille affirmation d'un nouveau champ scientifique immense et de telles perspectives multi- ou trans-disciplinaires, allant des sciences les plus « dures » aux sciences les plus « souples », et dont ce livre fournit un témoignage, sont des choses rares. Aux chercheurs d'en profiter et d'en faire profiter le plus grand nombre d'individus, le plus grand nombre de citoyens, le plus grand nombre d'hommes. La grande aventure de l'intelligence artificielle doit être et demeurer une grande aventure de l'intelligence humaine.