



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-02327501v1

HAL Id: hal-02327501

<https://hal.science/hal-02327501v1>

Submitted on 22 Oct 2019 (v1), 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

Les relations entre l'homme et l'outil, l'être et la machine, le vivant et le robot, le naturel et l'artificiel jouent un rôle essentiel dans l'évolution 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 est d'ailleurs à l'origine de l'histoire —, 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 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. 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.

Pourtant, les révolutions de l'intelligence artificielle n'en sont peut-être encore 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, la plupart des pratiques. En revanche, leurs divergences sont grandes quant au tour que prendront ces pratiques chamboulées. 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 »¹. Il s'agit d'un « ensemble de fonctions psychiques et psycho-physiologiques concourant à la

¹ V° « Intelligence », in *Trésor de la langue française*.

connaissance, à la compréhension de la nature des choses et de la signification des faits »². 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 »³. Sous cet angle, l'intelligence de l'intelligence artificielle est encore balbutiante et très incomplète, ce qui invite à œuvrer à cette intelligence de l'intelligence artificielle, 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 contributions de certains des meilleurs spécialistes de la question. Il existe un consensus 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'expliquer⁴. En ce sens, le besoin d'intelligence de l'intelligence artificielle serait réel. Jusqu'à la définition de l'intelligence artificielle, à base de critères et de 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 de quoi il s'agit. Elle 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 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 informatique qui effectue des actions ou exécute des tâches qui, il y a peu encore, étaient le propre des êtres vivants, humains ou animaux⁵. Cet outil repose 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. 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 ; elle est uniquement de façonner des outils capables de produire les résultats les plus pertinents possible, essentiellement grâce à de l'analyse statistique de données. Comme il s'agit d'une expression-valise, désignant des technologies disparates et variées, il serait logique de se référer aux « intelligences artificielles » (au pluriel).

« 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 d'imiter le

² *Ibid.*

³ *Ibid.*

⁴ 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.

⁵ 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 ».

comportement humain ou animal, tandis que l'« artificiel » se rapporte à l'usage d'ordinateurs et autres procédés électroniques. 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.

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 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. Un algorithme ne peut en effet décider en toute autonomie de ses buts ni concevoir de nouvelles représentations du monde qui le guideraient dans ses raisonnements. C'est pourquoi l'appellation « intelligence artificielle » est largement excessive. Il ne s'agit guère d'intelligence mais seulement de calculs (certes savants), de statistiques et d'opérations informatiques.

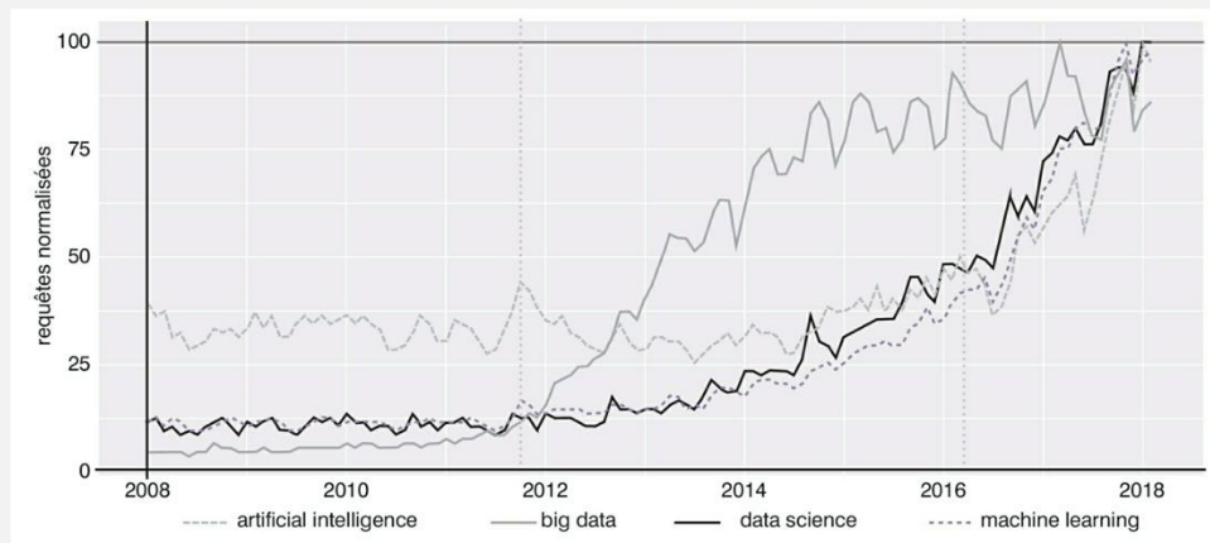
Parler d'« intelligence artificielle » est donc un abus de langage. « Intelligence artificielle » vient de l'anglais « *artificial intelligence* » désignant la gestion des données, le traitement de l'information, et non l'intelligence artificielle telle qu'on la comprend en français — qui se traduit plutôt par « *artificial cleverness* » et qui désigne le sens commun, l'empathie, l'arbitraire éclairé, la faculté de comprendre sa propre intelligence et d'en avoir conscience. En somme, en anglais, il n'est pas décevant 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, n'ont pas conscience d'eux-mêmes et ne possèdent aucun instinct de survie — ce qui constitue pourtant les marques de l'intelligence. Les hommes se caractérisent par l'esprit, par la liberté et par la capacité à s'ouvrir à l'inconnu. « Intelligence artificielle » est un anthropomorphisme. On humanise, on vitalise ce qui ne sont toujours que des objets animés, des automates, des supercalculateurs⁶. 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 qu'est conduire. L'expression « intelligence artificielle » (ou « IA ») s'étant partout imposée, tant dans les débats grand public que parmi les discussions de spécialistes, il est cependant difficile de désigner l'objet de ce livre par d'autres mots, sous peine de ne pas se faire comprendre.

La mode de l'« intelligence artificielle » et de son sigle « IA » est récente, comme en témoignent les statistiques de requêtes sur le moteur de recherche Google. Elle n'en est pas moins forte et, si l'on veut en parler, on ne peut y échapper.

⁶ 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* »⁷. Un nouveau domaine d'enseignement et de recherche scientifiques était né, dont l'objectif 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 opposer à raison cette dernière à 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. 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 parvenir — bien qu'on commence à appliquer le *machine learning* à

⁷ 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).

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. Mais celles-ci semblent difficilement pouvoir rejoindre ou dépasser l'intelligence naturelle, même à moyen ou long terme.

D'ailleurs, quand les chercheurs du Dartmouth College créèrent l'expression « *artificial intelligence* », il s'agissait avant tout d'une démarche marketing, presque un mensonge volontaire : recourir à une expression sensationnelle afin d'obtenir des crédits de recherche de la part de l'État américain⁸. L'idée d'intelligence s'attache à un fond profondément biologique et l'intelligence artificielle n'est donc guère intelligente. Mais cela ne l'empêche pas de rendre bien des services et de poser de nombreux défis à l'intelligence humaine, en premier lieu celui de son intelligence, c'est-à-dire de sa compréhension et de son explication.

L'intelligence de l'intelligence artificielle suppose de combiner quatre efforts cognitifs :

- accepter que l'IA n'est qu'un outil développé par et pour les hommes (I) ;
- s'intéresser aux rouages et au fonctionnement de la technologie (II) ;
- se concentrer sur les conséquences humaines et sociales de l'IA (III) ;
- donner la parole aux chercheurs et l'écouter (IV).

I. L'intelligence artificielle au service de l'intelligence humaine

Des robots en venant à défier et affronter les hommes est de l'ordre du 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 longtemps 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 fait⁹. Il n'y a et n'y aura pas plus de compétition entre l'homme et la machine qu'entre le pianiste et le piano ou entre la truie et le maçon. Et puis il serait difficile de dire qu'une espèce aussi proliférante que l'espèce humaine serait en train de s'effacer et de perdre la mainmise sur le monde, surtout au profit de ses propres technologies.

En 2014, l'astrophysicien Stephen Hawking affirmait, dans une interview à la BBC, que « le développement d'une intelligence artificielle complète pourrait provoquer la fin de l'humanité »¹⁰. Il s'inquiétait d'éventuels périls liés à l'avènement de robots qui surpasseraient des êtres humains limités

⁸ 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.

⁹ 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 ».

¹⁰ 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.

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 qui n'ait jamais pesé sur l'humanité —, entretiennent une psychose apocalyptique autour de la prise de pouvoir de machines toutes-puissantes créées par l'homme puis lui échappant¹¹. Il est difficile d'imaginer qu'un jour la machine pourrait 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. Et 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 — ; ainsi que celle selon laquelle il existerait un secret scientifique et technologique de l'intelligence artificielle qui resterait à découvrir et selon laquelle celui qui le découvrirait pourra 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 n'est pas celui de machines asservissant les humains mais celui d'humains asservissant d'autres humains grâce aux machines — un problème politique, juridique et économique. Là où il faut et faudra toujours de l'éthique, ce n'est pas dans l'intelligence artificielle mais dans l'intelligence humaine¹².

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 mais de permettre à l'humanité de progresser et d'entrer dans un nouvel âge. La finalité des projets d'intelligences artificielles devrait effectivement être de faciliter la vie des hommes et le bon fonctionnement de la société, de les protéger des dangers qui les guettent — la maladie, la corruption, la manipulation, l'aveuglement 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 fiables¹³. 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

¹¹ 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.

¹² 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.

¹³ 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.

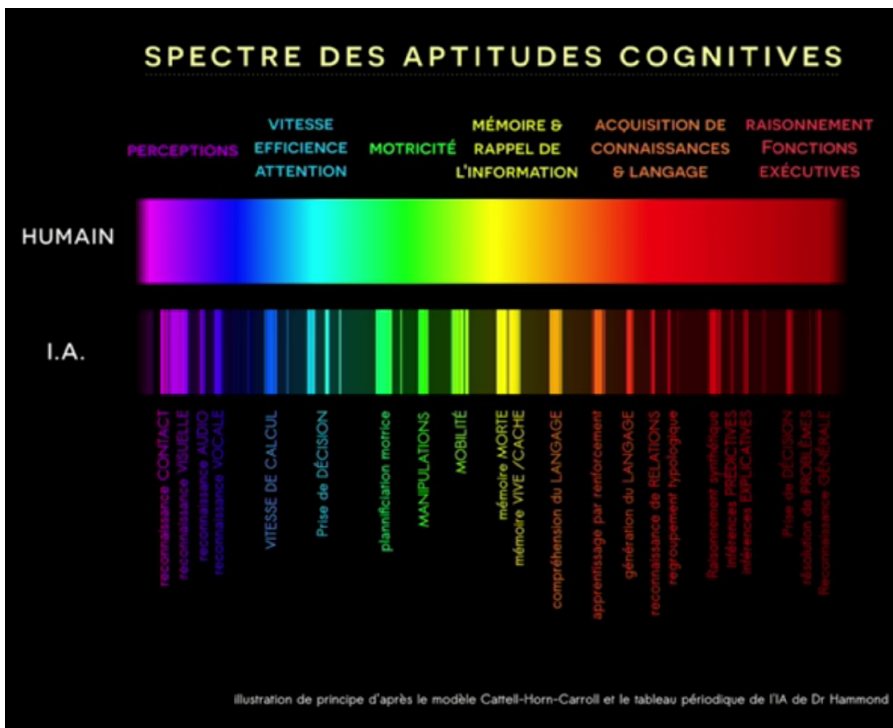
la première de toutes les menaces liées à l'IA¹⁴. Son utilisation comme outil de propagande ou de désinformation pourrait être désastreux. Elle est ainsi un moyen pour les personnes bien intentionnées comme pour les personnes mal intentionnés d'arriver plus facilement à leurs fins.

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 concurrencer¹⁵. 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.

Si les humains créent des formes d'intelligence artificielle, c'est afin de progresser, de rendre plus aisés nombre de leurs actes individuels et collectifs. 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 trier de grandes masses d'informations ou à réaliser des calculs hyper-complexes en quelques instants. L'IA est capable de faire des choses impossibles pour des humains, mais pas de faire des choses simples pour un homme, relevant 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 pouvant être très développées, quand les capacités cognitives de l'humain sont totales mais limitées par nature dans leurs performances :

¹⁴ 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.

¹⁵ 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'intelligence¹⁶. 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</p> <p>Capacité à penser, saisir et exprimer des idées. Exemples : écrivains, poètes, traducteurs, interprètes, journalistes, avocats, ...</p> <p>LES MOTS</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, accordeurs, ...</p>	
<p>Intelligence logico-mathématique</p> <p>123</p> <p>Capacité à raisonner, calculer, compter, résoudre des problèmes. Exemples : scientifiques, ingénieurs, savants, mathématiciens, comptables ...</p> <p>LA LOGIQUE</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 classer la nature. Exemples : explorateurs, botanistes, archéologues, zoologistes, fermiers, ...</p>	

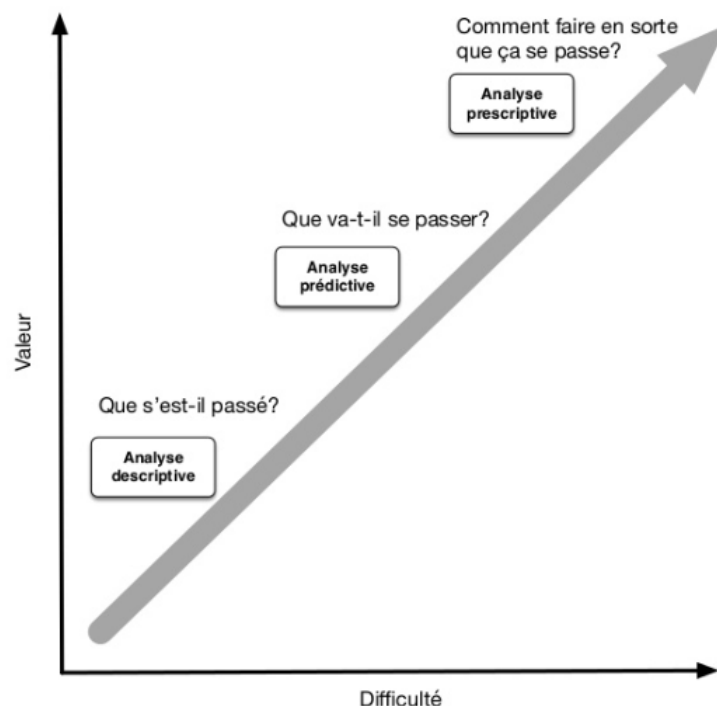
Illustrations : Céline Cidère / www.pour

¹⁶ H. Gardner, *Les formes de l'intelligence*, Odile Jacob, 1997.

C'est la symbiose de deux formes d'intelligence — ou association de l'intelligence et de ses réalisations les plus perfectionnées — qui permet le progrès des savoirs et des connaissances, ainsi que des idées. Et les chercheurs comptent au nombre de ceux qui en retirent les plus grands bénéfices. En même temps, et de façon de plus en plus forte, la frontière entre les tâches dévolues aux machines et celles qui restent réservées aux hommes ne cesse de se déplacer.

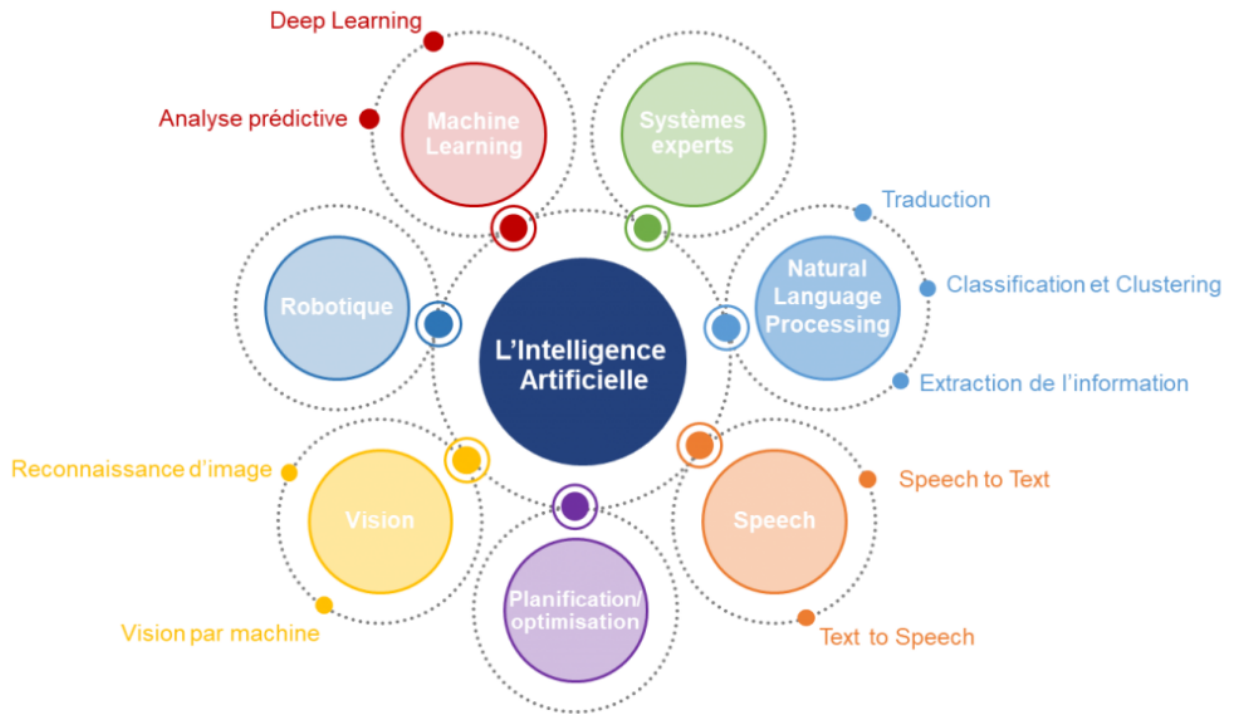
L'intelligence artificielle a vocation à servir l'intelligence humaine en éclairant et donc facilitant les prises de décisions. Pour ce faire, elle doit lui fournir des informations à haute valeur ajoutée, notamment avec le tri et le croisement d'informations et avec des analyses prédictives :

Ajouter de la valeur



CETIC, juin 2017

Les principaux domaines fonctionnels de l'IA témoignent en tout cas des innombrables réalisations auxquelles elle est susceptible de donner lieu. L'IA permet ainsi d'approfondir des domaines fonctionnels déjà connus, à l'instar des systèmes experts, de la planification/optimisation ou de la robotique, et elle permet de développer les nouveaux domaines que sont 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 cela a vocation à concerner de très nombreux secteurs :

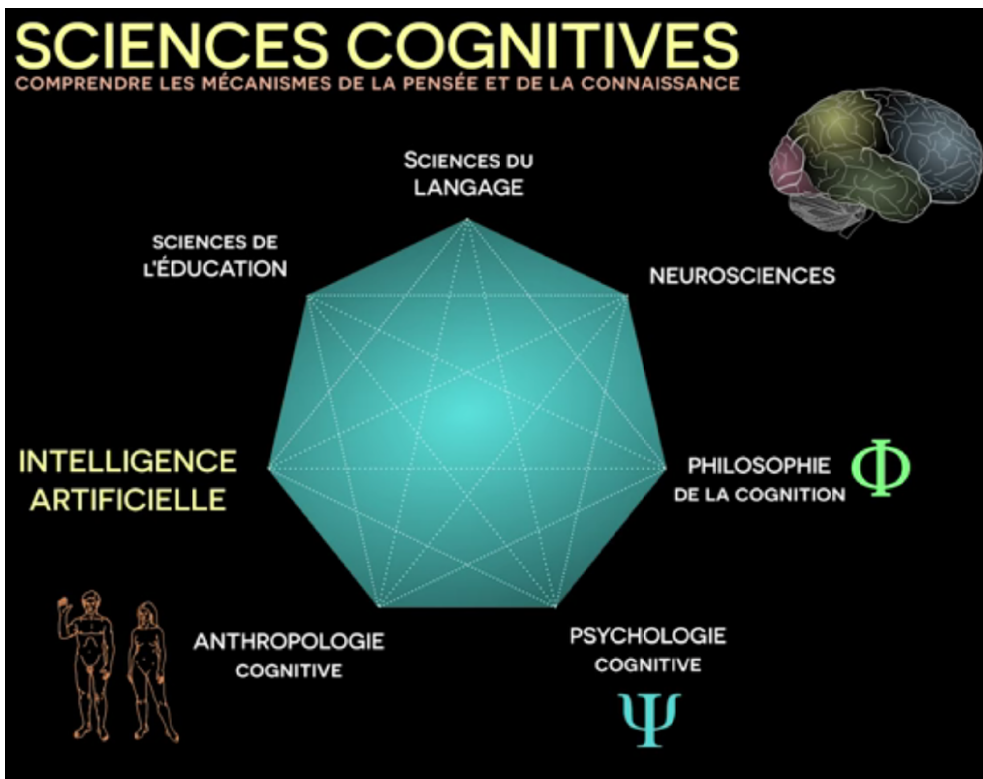
Utilisation des algorithmes et de l'intelligence artificielle

	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 : CNIL - Ethique et numérique

II. Au cœur des cerveaux artificiels : apprentissage automatique et apprentissage profond

« Intelligence artificielle » peut désigner un objet ou désigner une science. 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

Peut-être mieux vaut-il cependant n'utiliser « intelligence artificielle » qu'afin de désigner un objet et parler de « sciences de l'intelligence artificielle » au sujet des disciplines qui étudient cet objet — 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.

Si l'IA doit de préférence être conçue tel un objet, il serait bien sûr malvenu de chercher à disserter 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 comprendre, 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 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, mais pas nécessairement technoaddict.

Les origines de l'intelligence artificielle sont anciennes. L'IA s'inscrit 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 XVIe siècle, Léonard De Vinci conçut nombre d'automates imitant l'homme ou l'animal. Et, au XVIIe 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 a franchi un pallier considérable. Cette année là, pour la première fois, un réseau de neurones profond a remporté 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 a remis 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 — durant ce qu'on qualifie d'« hiver » de l'intelligence artificielle. 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 ») a permis à ces programmes d'enfin faire leurs preuves à grande échelle.

La communauté scientifique a alors perçu l'immense potentiel de cette technologie : en étant entraîné sur une base de données suffisamment consistante et pertinente, un programme informatique est capable d'établir lui-même des règles qui vont permettre d'interpréter d'autres données qui étaient 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 go¹⁷. Vingt ans après les 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 endurantes, 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 ».

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, mais, pour l'heure, c'est l'apprentissage grâce à des exemples qui domine très largement.

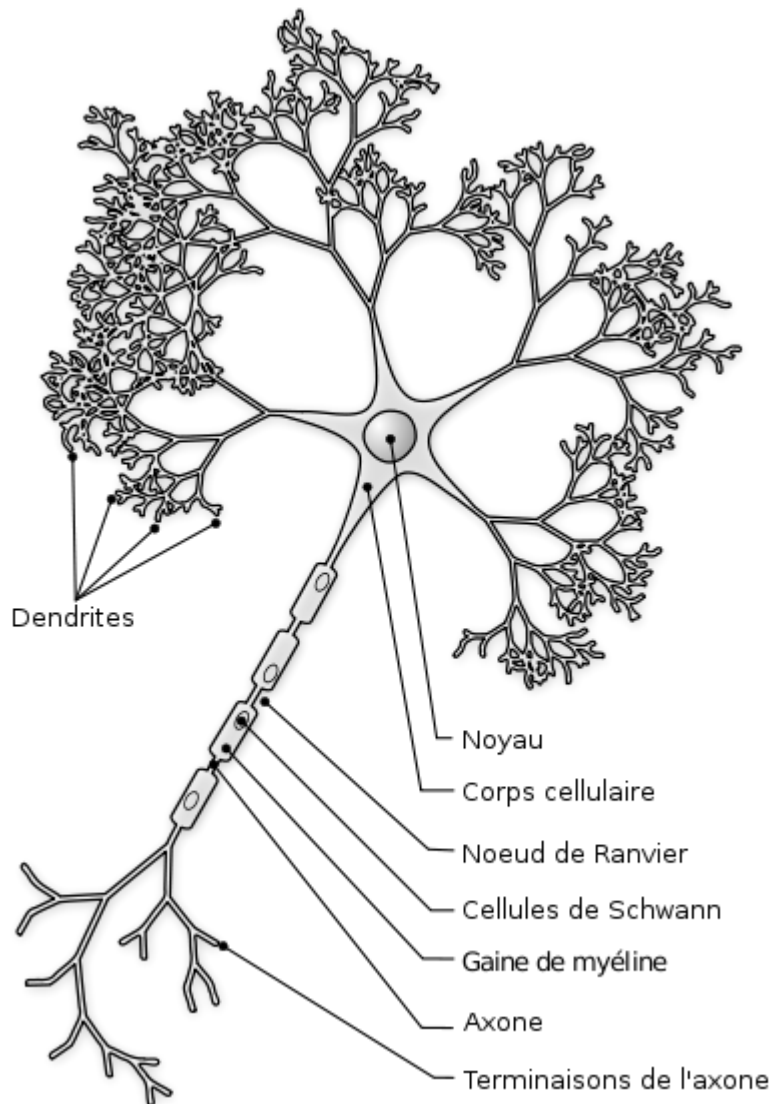
Pour entraîner une intelligence artificielle à reconnaître des chats sur des images contenant soit des chiens, soit des chats, on lui montre un grand nombre d'exemples d'images de chats et un grand nombre d'exemples d'images de chiens. Plus le nombre d'exemples est grand, plus la machine devient capable de reconnaître seule les images de chats car elle a en mémoire beaucoup de cas de figure possibles. C'est ce qu'on appelle l'apprentissage automatique (*machine learning*). Lorsqu'on parle d'intelligence artificielle aujourd'hui, c'est presque toujours 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.

¹⁷ 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.

Le *machine learning* est notamment utilisé en matière de recommandation (rechercher les consommateurs les plus proches d'un individu pour suggérer l'achat de produits), le *clustering* (définition de groupes typiques, par exemple d'utilisateurs d'un réseau social), la détection automatique de fraudes, anomalies ou erreurs de saisies. Le *deep learning* désigne quant à lui 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 complexes, comme la reconnaissance automatique de textes et 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).

Le cerveau est peut-être l'objet le plus complexe qu'on connaisse dans l'univers et son fonctionnement demeure 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 nerveux¹⁸.

¹⁸ 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.

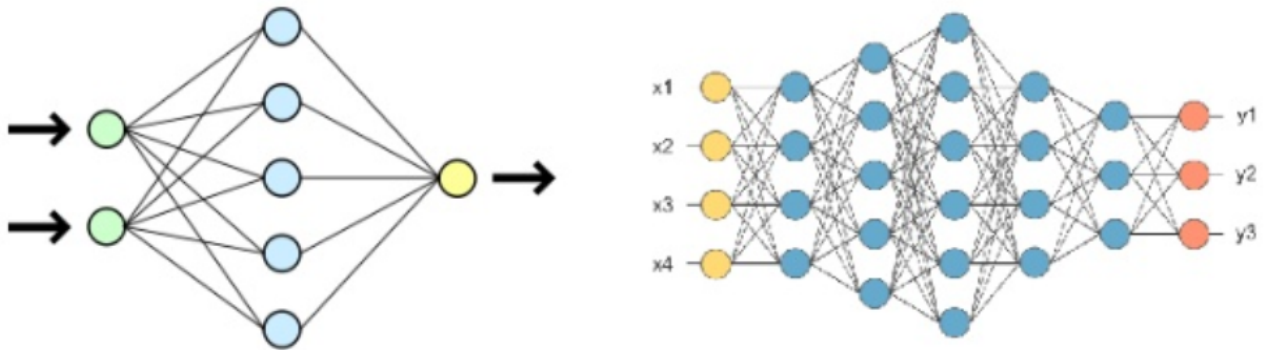


Nicolas Rougier, schéma biologique d'un neurone, 2007

Pour apprendre, un cerveau humain ajuste la force des interconnexions entre ses milliards de neurones. Personne ne le programme. Un réseau de neurones électroniques fonctionne de la même manière¹⁹. 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 « ceci est un chat ». 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

¹⁹ 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.

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.



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 possible²⁰. À ce niveau, cerveau artificiel et cerveau naturel fonctionnent très

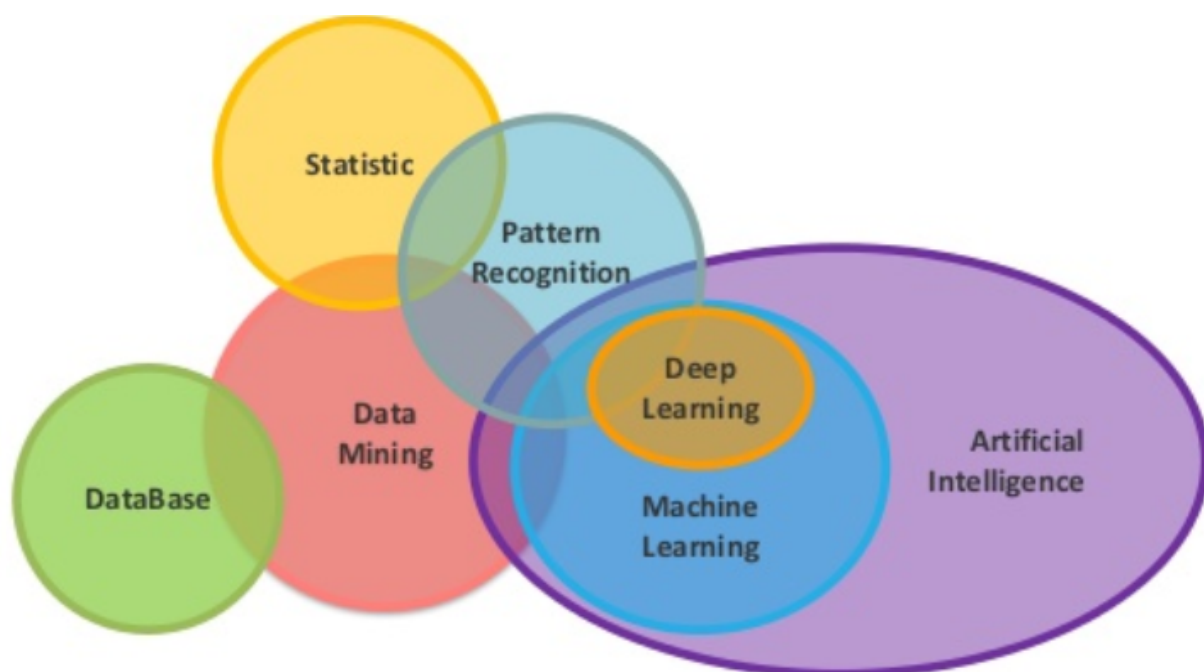
²⁰ Mais la quantité de données d'une seule image peut constituer une difficulté. Une photographie numérique contient plusieurs millions de pixels, chacun pouvant être considéré telle une donnée d'entrée pour le réseau. Le volume de données risque alors d'être beaucoup trop important pour être traité 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).

différemment. Les tsunamis de données sont anti-biologiques : un cerveau humain est incapable de traiter autant d'informations²¹. Au départ, un réseau de neurones est totalement incapable de distinguer un chat d'un chien, comme l'est le cerveau d'un bébé qui vient de naître. 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 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 données 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 d'exemples particuliers. On peut illustrer cela par le cas suivant : grâce à l'historique de nombreux acheteurs, un algorithme va déterminer que deux biens sont complémentaires parce que l'achat de l'un entraîne peu de temps après l'achat de l'autre. 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 et de créativité artificielle²².

L'intelligence artificielle a donc besoin du big data, 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 et l'apprentissage profond.

²¹ 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.

²² 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.



CETIC, juin 2017

L'intelligence artificielle apparaît sitôt qu'on parvient à obtenir, sur un objet donné, des résultats pertinents de façon automatique. Un thermomètre est ainsi une forme très primaire d'intelligence artificielle parvenant à saisir et à 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 ainsi 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'il soit un jour possible de dépasser le deuxième âge. Pour l'heure, la machine peut être plus intelligente que l'homme, voire beaucoup plus intelligente que lui, mais seulement concernant des tâches très spécifiques — comme elle peut être beaucoup plus forte physiquement que lui pour réaliser certains actes bien précis.

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 l'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 lorsqu'il n'y a pas de structure connue fournie à l'algorithme, seulement des exemples. À partir de ces derniers, la machine doit déterminer les structures cachées permettant de relier les exemples. 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. Cette technique 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 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. 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.

Les grands acteurs actuels de l'économie, puissances privées qui concurrencent les puissances publiques sur la scène internationale, se sont développés autour d'algorithmes, sortes de recettes de cuisine jalousement conservées, les ingrédients étant les gigantesques masses de données constitutives du *big data*. Mieux vaut avoir un mauvais algorithme avec beaucoup de données qu'un bon algorithme avec peu de données. Mais ces acteurs possèdent les deux. L'intelligence artificielle devient ainsi le noyau dur de beaucoup de secteurs de l'économie, de la recherche et de la vie. Ce sont les géants du numérique, GAFAM et BATX, qui possèdent les principaux réservoirs de données. C'est pourquoi ils s'imposent telles les principales puissances technologiques et économiques, mais aussi politiques et scientifiques. Et la majorité des innovations numériques et des technologies et applications de l'intelligence artificielle proviennent d'un écosystème public-privé unique : la Silicon Valley. Cela explique pourquoi 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 d'entraîner les IA²³. 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, Xiaomi²⁴), car produire une IA performante suppose un long apprentissage au moyen d'énormément d'informations, dont seuls ces géants du numérique disposent. 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 originales²⁵.

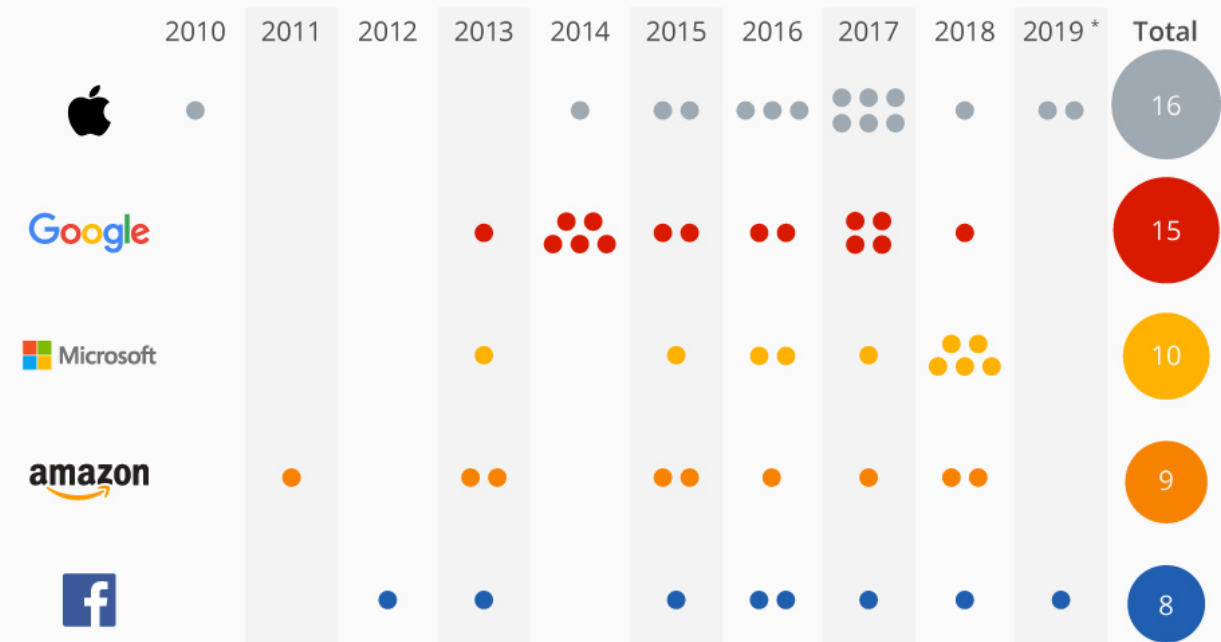
²³ France et Europe importent massivement les technologies IA américaines sur smartphones Android et iPhones.

²⁴ On peut y ajouter IBM, Samsung, Qualcomm, Intel, Philips, Siemens et Sony.

²⁵ 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.

Apple et Google en tête de course à l'intelligence artificielle

Nombre de start-ups IA rachetées par les GAFAM depuis 2010 *



* Les dernières données disponibles datent de mars 2019.

Sources : CB Insights, Statista Digital Market Outlook, rapports média

statista

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 refusant de financer cette recherche.

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 de cette technologie, 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 les chercheurs, 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 rapides 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 classement de publications à afficher en priorité sur les réseaux sociaux. Des problèmes moraux de taille 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 de l'algorithme.

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

III. L'homme et la société à l'épreuve de l'intelligence artificielle

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 confidentielles²⁶. 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 et les applications mobiles), des torrents d'informations sont en permanence livrés par chacun et stockés dans d'immenses serveurs informatiques.

Quant à la seconde dynamique qui contribue fortement à dessiner le futur, elle est liée au besoin de comprendre et interpréter ces gigantesques ensembles de données. C'est ici qu'interviennent les algorithmes, donnant aux ordinateurs des instructions mathématiques pour trier, traiter, agréger et représenter les informations. Ainsi de puissantes technologies statistiques et « prédictives » se développent-elles 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 parmi des ensembles gigantesques d'informations beaucoup mieux que l'intelligence humaine. Les classements, palmarès, compteurs, cartes, recommandations et notes diverses qui s'affichent en permanence sur les écrans des 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, 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

²⁶ 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.

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.

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. Ces données peuvent aussi être utilisées afin de créer de nouvelles formes de servitude. C'est pourquoi la politique et le droit doivent intervenir, à l'image de l'Union européenne et du Règlement général sur la protection des données (RGPD) applicable depuis le 25 mai 2018²⁷. 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 à quelques-uns. L'enjeu, les concernant, 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 grave obstacle 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. Lorsque la confidentialité des données relève de la protection de la vie privée en Europe, il apparaît suspect de vouloir protéger sa vie privée dans l'« Empire du Milieu ». En Chine, vouloir protéger ses données est perçu comme de l'égoïsme parce qu'on 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 grâce à l'IA. 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 accroître son *soft power*²⁸. Mais faudrait-il pour autant suivre cet exemple ?

La Chine ne souhaite cependant pas se couper du monde et, tandis que certains GAFAM s'établissent sur son sol afin de profiter de certains de ses talents en la matière, 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 de la déclaration 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, sont donc perçues comme collectivistes et utilitaristes, ce qui est en parfait accord avec la philosophie confucéenne — qui connaît d'ailleurs un fort renouveau.

²⁷ 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.

²⁸ 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.

Historiquement, un pays qui domine les autres puissances technologiquement et économiquement a toujours été impérialiste. Même si la Chine annonce travailler aux progrès de l'humanité²⁹, il est difficile de croire qu'elle n'a 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, avec le *soft power*, en s'imposant dans les esprits et dans les pensées. Tout en demeurant au fond de son Cantal natal, on peut devenir « sujet » de l'empire chinois en y envoyant ses données et en utilisant ses services.

Par ailleurs, 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*³⁰. *Data scientist* serait pourtant le « métier le plus sexy du XXIème siècle », selon la *Harvard Business Review*. Mais le World Economic Forum s'inquiète ainsi du fait que les travailleurs, notamment français, seraient très mal armés pour s'adapter et affronter le changement. L'intelligence artificielle pose donc le problème de l'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 universel³¹. 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, 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. S'il n'y a plus besoin des hommes pour les métiers basiques, parce que l'IA s'en chargera 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 transversales³², où il y a peu de données et de statistiques mais besoin de sens commun, de souplesse, d'adaptabilité, de créativité, de culture générale, de sens critique, d'empathie, d'affectivité, d'humanité, de capacité à prendre des décisions arbitraires et à innover, ainsi que, parfois, de force physique et de dextérité. L'IA est gratuite en valeur relative par rapport à

²⁹ Cf. Kai-Fu Lee, *IA. La plus grande mutation de l'histoire*, Les Arènes, 2019.

³⁰ 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.

³¹ 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.

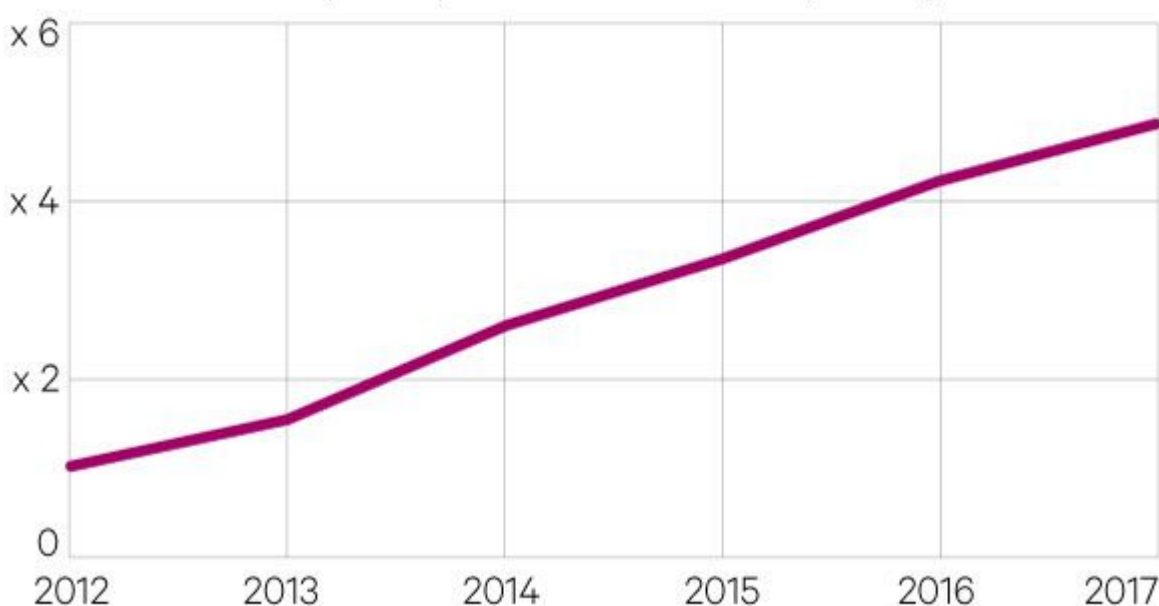
³² L'IA peut concurrencer un radiologue mais pas un médecin généraliste. Face à l'IA, le généraliste garde toute sa place, au contraire du spécialiste.

l'intelligence biologique³³, 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 prennent beaucoup de valeur. Ici, ce qui est 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 américaines³⁴ a en moyenne plus que quadruplé depuis 2012.

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

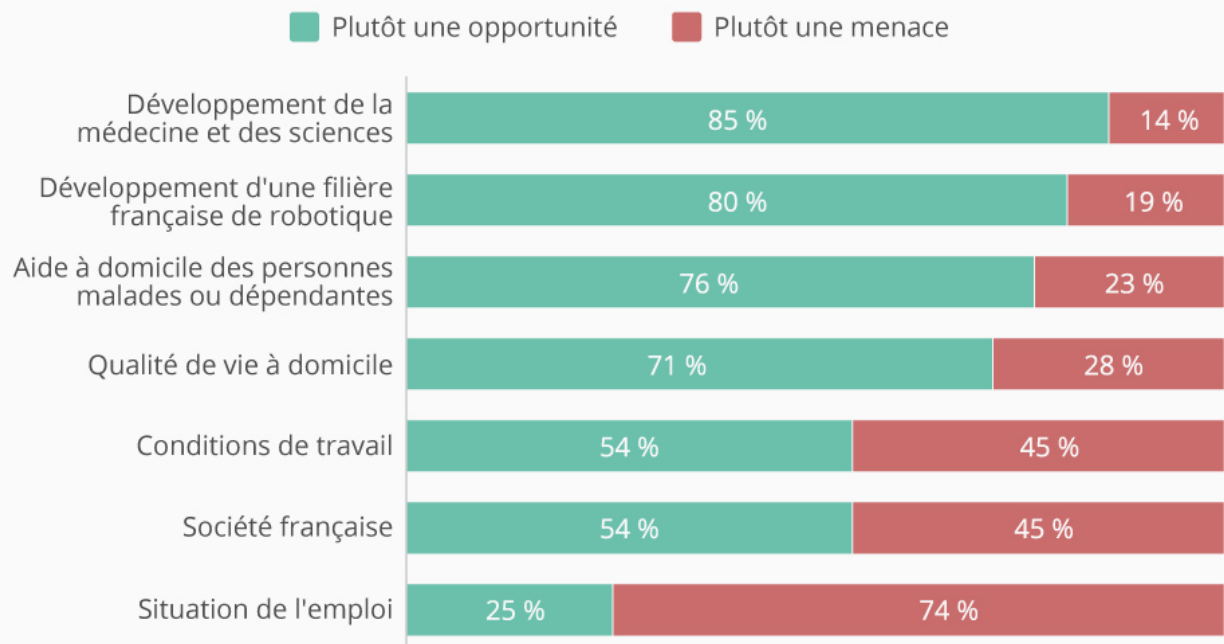
Ce sont logiquement les conséquences du déploiement des IA sur la situation de l'emploi qui suscitent le plus de craintes, tandis que, à l'inverse, ses répercussions en matière de développement de la médecine et des sciences génèrent le plus d'espoirs :

³³ 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, ce qui laisse, pour un temps, à l'abri des métiers tels que femme de ménage.

³⁴ Berkeley, University of Illinois Urbana-Champaign, Stanford, Carnegie Mellon et University of Washington à Seattle.

L'intelligence artificielle, chance ou menace ?

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



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

statista

Si l'optimisme semble l'emporter sur le pessimisme en matière de perspectives de l'intelligence artificielle, celle-ci n'est cependant pas le sujet intéressant le moins la « collapsologie », ce « discours sur ce qui s'effondre », en l'occurrence la civilisation industrielle du XIXe et du XXe siècles. Elle touche aussi les survivalistes, dont les adeptes, qui brandissent les menaces économique, climatique, civilisationnelle et technologique, se préparent à une très prochaine fin du monde. Selon les collapsologues, plus une société est inégalitaire, plus elle risque d'exploser. L'IA pourrait être le terreau d'une grave crise sociale. En tout cas pose-t-elle de redoutables questions éducatives, philosophiques et politiques. La démocratie pourrait ne pas survivre au creusement des fossés au sein de la population, notamment en termes d'intelligence, de culture et d'éducation. La « trumpséxisation » 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, 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 d'hommes 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 aristocratie de l'intelligence et un prolétariat désœuvré, démuné, abandonné, entre ceux qui voguent sur la vague des nouvelles technologies et les naufragés du numérique et de la société de la connaissance. 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 plutôt que plaintif, est-il plus intéressant.

L'IA pourrait être le prochain « cliquet malthusien », permettant à l'humanité de faire un bond en avant aussi bien d'un point de vue 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. D'ailleurs, 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 doit 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 Shangai 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 mort³⁵, coloniser mars, augmenter le cerveau et l'intelligence humaine) etc.

La tendance à chiffrer et enregistrer toutes sortes d'informations est l'objet de nombreuses critiques ; 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 traitent³⁶.

Les plus grandes surprises proviennent en réalité moins des nouvelles technologies que des usages sociaux ou économiques qui en sont faits, plus disruptifs encore. 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. Cette personnalisation accompagne l'individualisation dans les sociétés contemporaines, le creusement du fossé entre chacun et le collectif, le besoin croissant de liberté et d'autonomie — qui peuvent n'être qu'apparents. Avec un effet performatif fort, les algorithmes aident à prendre des décisions ; et, souvent, ils les prennent eux-mêmes. Comme l'invention du microscope ou celle du télescope ont permis de voir autrement le monde et l'univers, les capteurs numériques, toujours plus omniprésents, mesurent tout et prédisent tout. 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, dans une large mesure, l'avenir de leurs disciplines, du point de vue des objets comme des moyens d'étude.

³⁵ 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.

³⁶ 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.

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 ces nouvelles technologies seront les plus importants et les plus radicaux.

Le présent ouvrage en appelle donc 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 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.

IV. Un enjeu d'intelligence humaine : pour une science grand public de l'intelligence artificielle

Les contributions contenues dans ce livre, rédigées par les meilleurs spécialistes de chaque question, 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 philosophique et prospectif. Par ailleurs, comme y invite la collection Europe & Asie, quelques commentaires concerneront 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-Unis³⁷.

Une radiographie critique de l'intelligence artificielle et de ses conséquences est un enjeu démocratique et social fort. Dans tous les domaines, 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 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

³⁷ 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.

algorithmique. Pourtant, habités par un sentiment d'incompétence et d'éloignement, bien que les algorithmes soient partout et tout proches, on préfère souvent se désintéresser de la question, considérant qu'elle ne serait qu'affaire de spécialistes. Cet ouvrage souhaite ainsi 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.

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 expliquer au plus grand nombre le fonctionnement des algorithmes et des big data. 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 chacun³⁸.

Les algorithmes sont forcément politiques et même idéologiques, jamais neutres. On fabrique ces 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 ne sera 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.

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 que le plus grand nombre comprenne 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 renforcer le suivisme, le panurgisme, à limiter le libre-arbitre et la diversité. L'efficacité et le confort des intelligences artificielles omniprésentes dans les décisions et les conduites quotidiennes de chacun tendent à endormir le sens des responsabilités, l'autonomie et l'individualité³⁹. Or ne faut-il pas conserver, pour reprendre une expression de John Stuart Mill, un « droit à l'errance », une capacité à dévier, à délibérer intérieurement, à s'autogouverner, à se choisir, à être libre en somme⁴⁰ ? 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

³⁸ 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.

³⁹ Notamment, G. Koenig, *La fin de l'individu – Voyage d'un philosophe au pays de l'intelligence artificielle*, Éditions de L'Observatoire, 2019.

⁴⁰ 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.

l'intelligence humaine. Ce « droit à l'errance » est pourtant aussi un « droit à l'erreur », lequel est une condition du progrès. Le risque est celui de l'immobilisme⁴¹ — par exemple, la *smart city* serait une ville morte, sans charme et sans humanité. L'optimisation peut être source d'inertie.

C'est ce qui interpelle l'historien israélien Yuval Harari lorsqu'il demande si les applications industrielles actuelles de l'IA, qui en viennent à connaître les individus mieux que les individus eux-mêmes, ainsi que les « *nudges* » de plus en plus omniprésents dans l'environnement, ne conduisent pas à déléguer la capacité de choisir et décider à des machines, donc à s'y soumettre dans des pans entiers de la vie individuelle, sociale et professionnelle⁴². La mathématicienne et militante américaine Cathy O'Neil déplore en ce sens que des « armes de destruction mathématiques » 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 élections⁴³. Il serait donc impératif de prendre conscience et de se protéger des dérives des algorithmes, rendues possibles par la connexion permanente des individus et par la puissance de calcul toujours plus importante des ordinateurs. Si l'IA, désormais au cœur de toutes les industries et industrie en soi, creuse dangereusement les inégalités et fragilise la démocratie, comme l'observe Cathy O'Neil, alors elle ne saurait demeurer hors de contrôle, loin de toute régulation⁴⁴.

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 contrôle 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 indiquent 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 contrôle⁴⁵. Le résultat est cependant à peu près le même : l'avènement d'une société de la surveillance et de réduction des libertés.

Il faudrait n'abandonner à l'IA 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 mondiale⁴⁶, 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

⁴¹ 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.

⁴² Y. Harari, *Homo Deus – Une brève histoire de l'avenir*, Albin Michel, 2017.

⁴³ 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.

⁴⁴ 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.

⁴⁵ 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).

⁴⁶ Avec le besoin de prendre garde à ce que les grands laboratoires américains n'aspirent pas tous les meilleurs spécialistes européens.

importance est d'autant plus grande que, s'agissant d'un sujet déterminant pour l'avenir de l'humanité, il est nécessaire de pouvoir en débattre en toute indépendance, sans défendre des intérêts technologiques, politiques ou économiques. Ainsi le grand public peut-il accéder à des informations objectives et à des avis éclairés. En même temps, les scientifiques doivent aider les gouvernants à faire les bons choix, des choix structurants au moment où l'on prend des virages décisifs.

Il n'y a cependant pas de chemin allant 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 les algorithmes, aussi perfectionnés et efficaces soient-ils, ni par les chercheurs. 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 individu 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 prospère⁴⁷.

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 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 politique⁴⁸. 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 de crise de la rationalité. 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, d'économie de l'IA, mais en avance en matière de pensée de l'IA et d'éthique de l'IA ? Or, 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 ordinateur-espion muni de deux caméras, un micro et une ribambelle de programmes ultra sophistiqués de surveillance et de traçage. On propose, par exemple, de consacrer une propriété privée des données personnelles afin de permettre à chacun de redevenir maître de lui-même. Ce système aboutirait à monétiser les informations livrées par les particuliers aux grandes plateformes numériques.

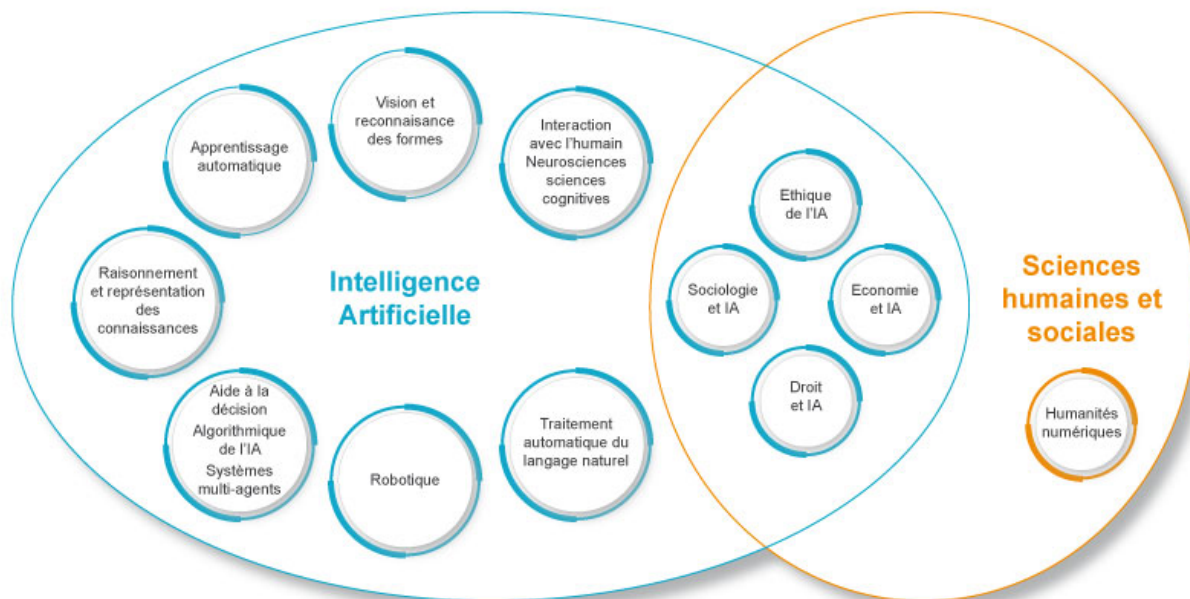
Le rapport entre le couple droit-politique et l'intelligence artificielle est d'autant plus important — et intéressant — que, si le droit et la politique saisissent l'IA (statut juridique, propriété intellectuelle,

⁴⁷ 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.

⁴⁸ 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 ».

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, civitechs, performativité des algorithmes etc.).

Alors que les spécialistes de la technologie, dans la Silicon Valley comme ailleurs, peuvent se laisser absorber par la seule question 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 ou la sociologie 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 diriger.



C2RP, septembre 2017

Et ces considérations deviendront sans doute toujours plus importantes avec l'avènement de l'homme augmenté et du transhumanisme. Pour l'heure, l'homme n'est que légèrement augmenté, avec un téléphone-ordinateur portable, une connexion 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 artificielle⁴⁹ pourrait s'analyser tel un progrès ou telle une régression, si ce n'est tel le début de la fin de l'humanité, tandis que « IA » serait le nom du nouveau dieu⁵⁰. « Homme augmenté », « intelligence augmentée », « *Brain editing* », « *neurohackers* »,

⁴⁹ 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.

⁵⁰ 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

« neuroprotecteurs », « neuro-intimité », métier de paramétreur d’implants intra-cérébraux, entrée du cerveau dans le monde de l’internet des objets, eugénisme technologique⁵¹... La grande rupture civilisationnelle est-elle proche ? En tout cas n’est-elle pas impossible. En pompiers pyromanes, certains sont tentés de fusionner avec le feu. Quelques garde-fous semblent alors utiles.

Bien sûr, toute morale est contingente et relative, variant dans le temps et dans l’espace, ainsi qu’entre les individus d’une même population. La question de l’IA et de la morale est donc difficile et même 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 réellement 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 ?

Il n’en faut et faudra pas moins réguler et non laisser faire les robots, les hommes augmentés et le nouvel « *homo deus* ». On pourrait craindre, par exemple, la prolifération d’une nouvelle forme de racisme : une discrimination et un rejet cognitifs, des plus intelligents et des augmentés envers les moins intelligents et les non augmentés. Et puis il faudrait prévenir la tentation des générations futures de tomber dans la datacratie, l’algocratie, le dataïsme et la datadépendance, c’est-à-dire l’enfermement dans un monde de chiffres, de statistiques et d’algorithmes déshumanisant et dénaturant. Peut-être mieux vaut-il conserver le hasard qui définit depuis toujours la vie et la marche du monde. La pantoprédicibilité et la dictée sociale des algorithmes doivent demeurer de simples cauchemars.

Les bioconservateurs pourraient-ils triompher et interdire ou supprimer les technologies ? Ou bien l’anarchie technologique pourrait-elle l’emporter, avec un développement tous azimuts uniquement motivé par la recherche de nouveaux marchés et de nouveaux monopoles ? Peut-être, comme souvent, la meilleure voie à suivre est-elle 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 intelligences artificielles 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 devant servir l’intérêt public.

Pour aborder sans détours et sans biais idéologiques les dilemmes d’aujourd’hui et changer 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, des articles aussi bien de mathématiciens que de spécialistes des sciences de l’information et de la communication ou des sciences humaines et sociales⁵². L’explosion de l’intelligence artificielle s’accompagne donc — et c’est heureux — d’une explosion des sciences de l’intelligence artificielle et, par suite, des

« divinité » basée sur une intelligence artificielle et qui est censée permettre le progrès de la société. Cela accrédite 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.

⁵¹ 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 ?

⁵² 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.

publications scientifiques spécialisées. Cela s'est produit en 2012-2013, au moment du printemps de l'apprentissage profond et des réseaux de neurones :

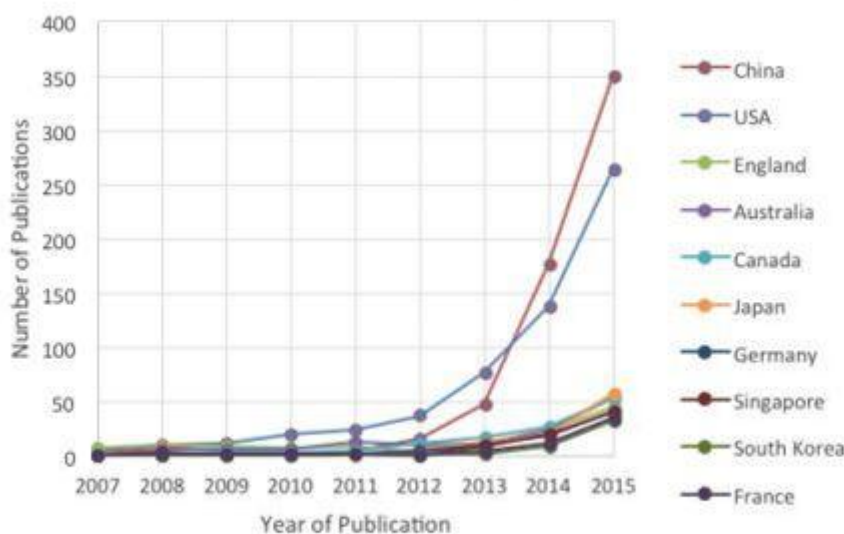


Figure 1: Journal articles mentioning "deep learning" or "deep neural network", by nation.⁶²

L'avenir est investi de toutes sortes de peurs, manipulé par le personnel politique à des fins électoralistes, tandis que le politiquement correct interdit de pouvoir 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'intelligence⁵³. 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'IA⁵⁴ 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 NBIC⁵⁵ progressent très vite, quand les institutions, notamment politiques et éducatives, ne changent guère, apparaissant parfois même sclérosées.

Dans ce contexte, les 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. Or il faut 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.

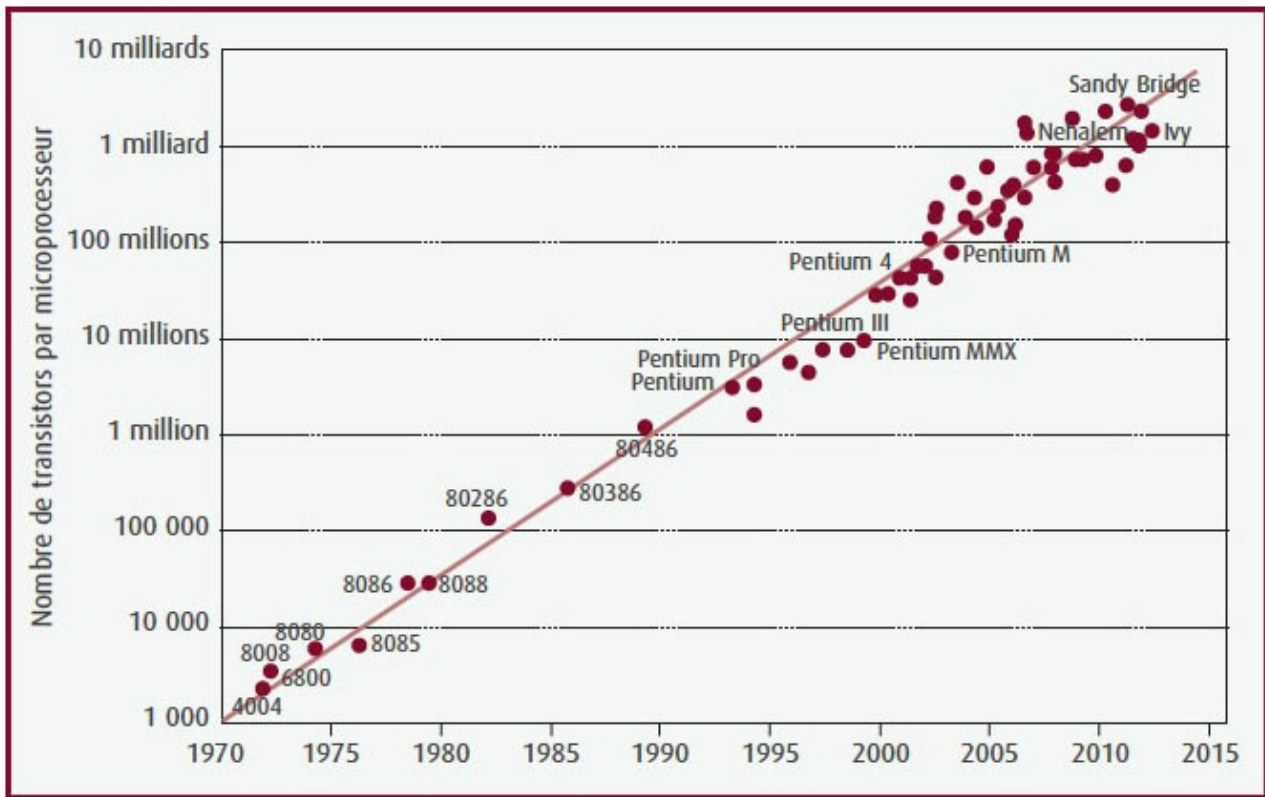
Cependant, le vrai pouvoir tend à échapper aux pouvoirs publics et à migrer vers les technologies, les infrastructures et donc ceux qui les possèdent. Et les progrès exponentiels des technologies NBIC les

⁵³ 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.

⁵⁴ 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.

⁵⁵ Les technologies NBIC sont les nanotechnologies, les biotechnologies, l'informatique et les sciences cognitives.

rendent largement imprévisibles. 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 avec la calculatrice mécanique de Blaise Pascal. La loi de Moore pourrait-elle se prolonger encore longtemps ?



Un autre frein à une meilleure appréhension du futur, de ses potentialités et de ses risques, est qu'aujourd'hui les mêmes canaux de communication diffusent les connaissances, l'information, la désinformation, les commentaires, les croyances et les opinions. Même les chercheurs s'expriment sur les réseaux sociaux. Ces contenus, qui n'ont pas le même statut mais qui se confondent dans les mêmes contenants, 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 se renforce. 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.

Sergey Brin, co-fondateur de Google, promet : « Nous allons créer des machines qui raisonnent, pensent et font les choses mieux que les êtres humains ». Si seul le temps permettra de vérifier la pertinence de pareille prophétie, 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 remodelera 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. Ce qui est 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, 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 humaine.