

# Intelligence artificielle, nouveau monde

written by Jean-Gabriel Ganascia | 16 janvier 2017

**Interview de Jean-Gabriel Ganascia**

Propos recueillis par [Françoise Laugée](#)

**Deep learning, réseaux de neurones, apprentissage supervisé, apprentissage non supervisé, apprentissage par renforcement...  
En quoi ces termes renvoient-ils à ce que l'on nomme intelligence artificielle ?**

Pour bien comprendre la signification de ces termes, il faut se référer à l'histoire de leur apparition qui advint très tôt, au XX<sup>e</sup> siècle, dans les années 1940 et 1950. Plus précisément, tout débute en 1943, avant même la construction des premiers ordinateurs électroniques, alors que l'on commence seulement à fabriquer des calculateurs électromécaniques, c'est-à-dire des machines capables de réaliser des opérations arithmétiques très rapidement, au moyen de relais téléphoniques. L'idée, assez naturelle somme toute, de dresser un parallèle entre ces machines et nos cerveaux, traverse les pensées d'un brillant mathématicien, Walter Pitts, qui, âgé d'à peine vingt ans à l'époque, écrit avec un neurologue, Warren McCulloch, un article intitulé « Un calculateur logique des idées immanentes dans l'activité nerveuse ».

L'IDÉE, ASSEZ NATURELLE SOMME TOUTE, DE DRESSER UN PARALLÈLE  
ENTRE CES MACHINES ET NOS CERVEAUX

Il y démontre qu'en organisant des automates élémentaires en trois couches, on peut réaliser n'importe quelle fonction logique. Ces réseaux d'automates dressent une analogie grossière avec ce que l'on connaît à l'époque du tissu de nos cerveaux, en ce qu'ils constituent des entités analogues aux cellules du cerveau, les neurones, et que ces entités se connectent entre elles par des liaisons pondérées semblables aux synapses correspondant aux liaisons physico-chimiques qui relient les neurones de nos cerveaux. Pour souligner le parallèle entre ces dispositifs et nos cerveaux, on les appelle aussi des réseaux de neurones formels.

Avec eux, il existe donc un pont entre ce que les neurosciences disent du cerveau, la logique, c'est-à-dire les lois de la pensée, et

l'ingénierie. Ce résultat majeur va donner naissance à une discipline nouvelle, la cybernétique, qui étudie les lois de la régulation de phénomènes complexes, telles celles qui régissent le vivant, par exemple les hormones ou les neurones, et la société.

Toutefois, même si de tels réseaux de neurones formels organisés en trois couches permettent de réaliser n'importe quelle fonction logique, il convient de configurer les liaisons synaptiques entre les neurones formels, autrement dit d'associer à chacune de ces liaisons un nombre, ce qui est extrêmement fastidieux, voire inextricable manuellement. On recherche donc, dès le début des années 1950, des procédures pour établir les pondérations des liaisons entre les synapses formelles. On imagine alors de donner à une machine des exemples étiquetés afin qu'elle ajuste automatiquement les poids des synapses formelles afin de retrouver automatiquement les étiquettes des exemples : c'est ce qu'on appelle l'apprentissage machine. Plus précisément, cet apprentissage est dit supervisé car un « professeur » donne des exemples avec leurs étiquettes.

À titre d'illustration, si l'on donne des formes géométriques comme des losanges, des polygones, des cercles ou des carrés, on indique à la machine que ce sont des losanges, des carrés, des pentagones, des cercles, des ellipses etc. Et on espère quelle sera en mesure de distinguer automatiquement cercles, ellipses, pentagones, quadrilatères si on lui donne suffisamment d'exemples ainsi étiquetés. Cela s'oppose à l'apprentissage dit non supervisé où les exemples sont donnés à la machine sans que personne indique la catégorie dont ils relèvent. On espère alors trouver des procédures capables de regrouper ensemble des exemples similaires et de dégager automatiquement des caractéristiques de ces groupes, comme les hommes le font dans la nature, par exemple lorsqu'ils distinguent, parmi les figures géométriques planes, les coniques et les polygones, sans que l'on ait évoqué ces notions auparavant. C'est là, à l'évidence, une tâche extrêmement délicate, bien plus que l'apprentissage supervisé.

Cependant, dans les années 1940 et au début des années 1950, même l'apprentissage supervisé apparaît très difficile, et les nombreuses tentatives qui ont été faites se sont conclues par des échecs. Heureusement, à la fin des années 1950, trois événements vinrent faire évoluer la situation, sans toutefois la débloquer totalement.

En 1956, deux jeunes mathématiciens, John McCarthy et Marvin Minsky organisent une école d'été financée par la fondation Rockefeller pour promouvoir une nouvelle discipline qui vise à décomposer l'intelligence en modules si élémentaires qu'une machine pourrait aisément les simuler. Selon eux, cela permettait de mieux comprendre les différentes facettes de l'intelligence et, éventuellement, de rendre de grands services à l'humanité en automatisant de nombreuses tâches qui requièrent de

l'intelligence. Ils décident de nommer cette discipline « intelligence artificielle ». Cela recouvre, entre autres, le perfectionnement des ordinateurs, la conception de nouveaux langages de programmation, la simulation du dialogue en langage naturel, ainsi que l'étude des réseaux de neurones formels, de l'apprentissage et de la créativité des machines, etc.

## LES NOTIONS ESSENTIELLES DE L'INTELLIGENCE ARTIFICIELLE ET DE L'APPRENTISSAGE EXISTENT DEPUIS PLUS D'UN DEMI-SIÈCLE

Un an plus tard, en 1957, un psychologue américain, Frank Rosenblatt, met au point un algorithme d'apprentissage pour des réseaux de neurones formels à deux couches qu'il appelle des perceptrons, car ils reproduisent selon lui les capacités de perception des rétines. Enfin en 1959, un ingénieur, Arthur Samuel, dote un programme informatique capable de jouer au jeu de dames de capacités d'apprentissage par renforcement. Il ne s'agissait là ni d'apprentissage supervisé ni d'apprentissage non supervisé, mais de permettre à la machine d'améliorer progressivement ses performances grâce à l'expérience.

Bref, les notions essentielles de l'intelligence artificielle et de l'apprentissage, qu'il soit supervisé, non supervisé ou par renforcement, existent depuis plus d'un demi-siècle. On peut donc se demander pourquoi il a fallu attendre autant de temps avant que le grand public ne mesure l'importance qu'elles ont pour la société.

Cela tient à trois facteurs : deux facteurs techniques liés, l'un à la puissance de calcul des machines et l'autre à leur capacité à engranger de grandes quantités de données, en particulier grâce au web ; ainsi qu'à un facteur d'ordre scientifique dû à l'amélioration des algorithmes d'apprentissage. En effet, la procédure d'apprentissage des perceptrons ne marchait que pour des réseaux de neurones formels à deux couches. Or, si Walter Pitts avait bien montré que les réseaux de neurones à trois couches pouvaient réaliser n'importe quelle fonction logique, il n'en va pas de même pour les réseaux à deux couches, tant s'en faut.

En 1969, Marvin Minsky démontra que la procédure d'apprentissage décrite par Frank Rosenblatt n'apprend que des fonctions très simples, dites linéairement séparables. Il a fallu attendre le début des années 1980 pour que des mathématiciens conçoivent une procédure d'apprentissage capable d'apprendre sur des réseaux de neurones à plusieurs couches. En termes techniques, on appelle cette procédure la « rétropropagation du gradient ». Quelques années plus tard, d'autres mathématiciens cherchèrent à trouver les fondements théoriques de cet apprentissage.

Cela les conduisit à développer d'autres techniques d'apprentissage comme les machines à vecteurs de support (*Support Vector Machine*), qui furent bien souvent utilisées dans les années 1990 et au début des années 2000. Puis, à partir des années 2010, on se remit à faire de l'apprentissage avec des réseaux de neurones dits convolutifs qui font apparaître un très grand nombre de connexions, de l'ordre de plusieurs dizaines de milliers, voire d'un million, et beaucoup de couches – entre 10 et 15 – dont certaines restent « figées » en ce sens que les poids synaptiques y demeurent fixes, tandis que d'autres évoluent par apprentissage. En raison de cette multiplicité de couches, on caractérise ces techniques comme de l'apprentissage profond (*Deep Learning*).

CES GRANDES COMPAGNIES RIVALISENT PLUS SOUVENT LES UNES CONTRE  
LES AUTRES QU'ELLES NE CONSPIRENT TOUTES ENSEMBLE CONTRE LE  
RESTE DU MONDE

La comparaison des capacités d'apprentissage du *Deep Learning* avec celles des autres techniques d'apprentissage supervisé comparables sur des tâches de reconnaissance d'images, a montré que les techniques de l'apprentissage profond apprennent de façon efficace sur de grandes quantités d'exemples tout en surpassant les performances des autres techniques. C'est la raison pour laquelle elles se sont imposées ces dernières années. Cela étant, bien qu'elles apparaissent neuves, les techniques d'apprentissage profond reprennent les principes anciens d'apprentissage par rétropropagation du gradient, introduits dans les années 1980, qui, eux-mêmes, reprennent les principes d'apprentissage des perceptrons qu'ils généralisent.

**La recherche en intelligence artificielle est menée par les GAFAs. Pourquoi cette ruée récente et accélérée vers l'intelligence artificielle ? Quelle stratégie les entreprises américaines poursuivent-elles ?**

**Alors que la bataille technologique est lancée, pourquoi les géants internet choisissent-ils de promouvoir leurs technologies (TensorFlow de Google, CNTK de Microsoft, SystemML d'IBM...) en *open source* ?**

Le sigle GAFAs signifie « Google, Amazon, Facebook, Apple ». Dans l'idée de ceux qui l'emploient, il désigne, de façon générique, les grands acteurs de l'internet dont le pouvoir s'étend au rythme de la numérisation de la société, si bien que l'on parle parfois de GAFAM en ajoutant un M qui renvoie à la société Microsoft. Et, pour compléter le tableau, on pourrait aussi y adjoindre la société Tesla et même – pourquoi pas ? – IBM et bien d'autres encore qui jouent un rôle éminent dans les industries du numérique et dont on craint qu'ils ne se joignent

eux aussi à une sorte de complot planétaire des industriels de l'hypermodernité. Il faut toutefois savoir que le terme GAFa est surtout utilisé en France et qu'il demanderait à être discuté plus à fond, car un examen attentif des relations entre ces grandes compagnies montre qu'elles rivalisent plus souvent les unes contre les autres qu'elles ne conspirent toutes ensemble contre le reste du monde.

Cependant, indépendamment de leur volonté de puissance, réelle ou supposée, et de leurs rivalités, elles recourent bien souvent à l'intelligence artificielle qui joue un rôle important dans leurs stratégies. De multiples déclarations le confirment. On constate aussi des investissements importants qui se traduisent par l'ouverture de centres de recherche et par l'emploi d'un grand nombre d'ingénieurs spécialisés en intelligence artificielle. Parallèlement, on note, chose étrange, une aspiration à rendre publiques les technologies mettant en œuvre les principes d'intelligence artificielle. C'est ainsi que Google a annoncé en novembre 2015 qu'elle mettait en accès libre le système TensorFlow qui comprend des technologies d'apprentissage machine recourant à de l'apprentissage profond (*Deep Learning*) utilisées pour l'analyse d'images, pour la reconnaissance de la parole ou pour la conception de logiciels de questions-réponses. Un mois plus tard, Facebook dévoilait son serveur dédié à l'intelligence artificielle, Big Sur, et le mettait aussi en libre accès. Quant à Elon Musk, il fonda en décembre 2015 la société OpenAI destinée à mettre aussi en libre accès tous les logiciels d'intelligence artificielle.

Nous pouvons donc faire deux constats qui paraissent à certains égards contradictoires : d'un côté, l'intelligence artificielle apparaît cruciale à tous ces acteurs et, d'un autre côté, ces entreprises se proposent de partager gratuitement les techniques d'intelligence artificielle qu'elles développent. Ces deux attitudes semblent de prime abord antinomique car, si l'intelligence artificielle est stratégique, cela signifie qu'elle donne un avantage compétitif sur les autres, et donc qu'il serait préférable de ne pas dévoiler son savoir-faire en la matière.

Pour bien comprendre ce qui est en jeu ici et montrer en quoi il n'y a rien de contradictoire, en dépit des apparences, il faut d'abord expliquer pourquoi l'intelligence artificielle devient un enjeu majeur, puis montrer les facteurs clés du développement des techniques d'intelligence artificielle.

**UNE ENTREPRISE DOIT EXPLOITER TOUTES LES SOURCES D'INFORMATION  
QUI PERMETTENT D'ÉVALUER ET DE COMPRENDRE L'OPINION DES  
CONSOMMATEURS**

L'importance de l'intelligence artificielle tient à l'économie spécifique de la Toile qui s'est mise en place à partir de 2004, avec l'avènement du web 2.0. Pour saisir cette spécificité et l'apport de l'intelligence artificielle, rappelons qu'au début du développement de la toile, dans le courant des années 1990, beaucoup pensèrent que l'économie du web se développerait sur le modèle classique des révolutions industrielles du passé, à savoir que les premiers arrivants sur les nouveaux marchés acquerraient rapidement des situations de domination quasi monopolistique. Cela explique la bulle spéculative de la fin de années 1990, lorsque les investisseurs misaient sur tous les marchés en germe susceptibles de se développer, sans examiner le détail des technologies mises en œuvre ; cela explique aussi l'éclatement de cette bulle, au tournant des années 2000, car elle tenait à une évaluation médiocre du potentiel réel des nouvelles entreprises et surtout à une mauvaise analyse de la structure du tissu économique de l'internet.

Quelques années plus tard, en 2004, un certain nombre d'industriels qui, en dépit de la crise, croyaient toujours dans la viabilité économique du web, essayèrent de comprendre les clés du succès de grandes compagnies de l'internet comme Amazon et Google. Ils constatèrent que, dans le monde numérique, la réputation se défait aussi vite qu'elle se fait et qu'il n'y a pas de situation définitivement établie. En conséquence, pour se maintenir, une entreprise doit exploiter toutes les sources d'information qui permettent d'évaluer et de comprendre l'opinion des consommateurs afin de remédier au plus tôt à d'éventuelles insatisfactions, d'améliorer ses produits en fonction de la demande et de répondre à des rumeurs malveillantes.

Cela conduisit à la mise en place du web 2.0, c'est-à-dire du web participatif où l'on implique l'utilisateur dans la conception et, surtout, où l'on récupère ses appréciations et ses réactions. Toutes ces informations constituent d'immenses bases de données qui traduisent les goûts, les opinions et les critiques des utilisateurs. Il convient d'être en mesure de les exploiter pour améliorer la qualité des produits, les configurer au regard des besoins de chacun et définir la stratégie des entreprises. Dans ce but, l'intelligence artificielle et, plus particulièrement, l'apprentissage machine jouent un rôle central ; en effet, ces techniques permettent d'extraire des connaissances à partir de très grandes masses de données, ce que l'on appelle les *big data*. Bref, c'est pour interpréter ces immenses quantités d'information recueillies sur le web que les grands acteurs de l'internet manifestent le besoin d'intelligence artificielle et d'apprentissage machine.

Ces techniques sont connues depuis longtemps ; elles ne constituent pas, en tant que telles, des nouveautés. Ce qui importe, ce sont d'une part les « tours de main » et les astuces nécessaires à leur mise en œuvre

dans différents contextes et, d'autre part, les masses de données sur lesquelles les algorithmes d'apprentissage fonctionneront. Or, en rendant publics les programmes informatiques qui mettent en œuvre les algorithmes d'apprentissage machine, les grandes sociétés ne dévoilent pas tout leur savoir-faire. En revanche, en contrepartie de ce qui apparaît aux yeux du public comme un don généreux, elles espèrent s'attirer la sympathie de tous, améliorer leur réputation et séduire des développeurs d'applications qui renverront des retours d'usage susceptibles d'aider à améliorer ces algorithmes d'apprentissage. Enfin, en invitant les utilisateurs à lancer les procédures d'apprentissage sur leurs propres machines, les grands acteurs de l'internet pourront aussi récupérer des données précieuses.

LA SOURCE DE LA RICHESSE NE TIENT NI AUX ALGORITHMES NI À LEUR  
PROGRAMMATION MAIS AUX MASSES DE DONNÉES SUR LESQUELLES ILS  
SONT MIS EN ŒUVRE

À cet égard, il faut comprendre qu'aujourd'hui la source de la richesse dans le monde numérique ne tient ni aux algorithmes de traitement de l'information, que l'on connaît depuis longtemps, ni à leur programmation informatique, mais aux masses de données sur lesquelles ils sont mis en œuvre. Ce sont ces données que Google accroît chaque fois que vous soumettez une nouvelle requête à son moteur de recherche ou que Facebook engrange à mesure que les réseaux d'amis s'accroissent, ou encore qu'Apple accumule lorsque vous utilisez ses produits. Or, ces masses de données demeurent privées... Et, en dépit du grand mouvement contemporain d'*Open Science* qui vise à partager les résultats de la recherche publique, elles demeurent la propriété exclusive des grandes sociétés de l'internet qui n'ont jamais sérieusement évoqué leur mise à la disposition du grand public...

**De la voiture autonome à l'habitat intelligent, de l'usine connectée au diagnostic médical, les applications de l'intelligence artificielle vont se multiplier dans la vie quotidienne. Quels seront, selon vous, les principaux défis à relever, notamment en termes de sécurité et en termes d'éthique ?**

L'intelligence artificielle aide à concevoir des agents autonomes, c'est-à-dire des entités qui, à l'instar d'une âme rationnelle, au sens aristotélicien, sont capables de percevoir leur environnement, puis de raisonner afin de décider des actions qu'elles doivent accomplir pour satisfaire leurs buts, et enfin d'agir effectivement. Précisons que la notion d'« autonomes » s'entend ici au sens technique. Cela signifie que s'enchaînent des processus physiques qui vont de la prise d'information à l'action en passant par la décision, sans qu'aucun homme intervienne.

Bien évidemment, on doit distinguer cette autonomie, entendue au sens technique, du concept d'autonomie morale formulé au XVIII<sup>e</sup> siècle par les philosophes des Lumières. Plus précisément, ces entités « autonomes » enregistrent de l'information à l'aide de caméras, de microphones et de toutes sortes de capteurs, puis elles interprètent toute cette information de façon à construire une représentation de leur environnement.

Ces dernières années, le perfectionnement des techniques d'apprentissage machine a permis de faire des progrès considérables dans l'analyse de l'information issue des capteurs physiques. Cela accroît d'autant l'autonomie des agents. La deuxième phase tient à la simulation du raisonnement. Là encore, les récentes évolutions de l'intelligence artificielle ont grandement amélioré les performances des procédures de décision automatique. En conséquence, nous voyons se déployer des agents autonomes de plus en plus élaborés dans beaucoup de secteurs, par exemple dans les industries manufacturières, avec les robots de fabrication, sur la route, avec les voitures autonomes, dans les maisons, avec les aspirateurs ou les tondeuses à gazon, à l'hôpital, avec les agents de surveillance des patients, etc.

Les défis sociétaux consécutifs à ces déploiements tiennent à l'autonomie de ces agents qui soulèvent trois problèmes de fond. Le premier est lié à la responsabilité engagée en cas d'accident : est-ce celle du propriétaire, de l'utilisateur ou du constructeur ? *A priori*, le propriétaire ne devrait pas être inquiet ; il n'en va pas de même pour l'utilisateur et pour le fabricant. En effet, imaginons une voiture autonome qui causerait, par son comportement, des victimes. Cela peut tenir à une utilisation non conforme ou à un défaut de fabrication ; seule une enquête pourra le déterminer et, dans tous les cas, il faudra adapter la législation pour prendre en considération tous les facteurs d'imputation de la responsabilité.

LA LOI DE MOORE N'A DE FONDEMENT QU'EMPIRIQUE ET CE N'EST EN  
AUCUN CAS UNE LOI DE LA NATURE

D'après moi, pour régler ces questions juridiques, on devrait introduire des formes de certification analogues à celles qui existent aujourd'hui dans l'industrie du médicament avec les autorisations de mise sur le marché. Cela conduira très certainement à préciser, dans chaque cas, les conditions d'utilisation dans lesquelles les constructeurs s'engageront à assumer leur responsabilité en cas de manquement. En revanche, les projets d'attribution de droits aux robots, dont certains avocats se font les chantres, m'apparaissent relever de la fantaisie. En effet, même s'il ne s'agit là que d'une fiction juridique, comme l'est le droit



moral des sociétés, ce droit n'aidera en rien à résoudre les problèmes d'attribution de responsabilités.

Le deuxième point délicat provient des applications critiques, en particulier des applications dans le domaine militaire ou, dans une moindre mesure, dans le domaine de la santé et des transports, en particulier des véhicules autonomes. En effet, dans ces cas-là, le robot est censé faire des choix qui conduiront à mettre en danger, voire à tuer sciemment une personne. Jusqu'à quel point peut-on froidement accepter ce type de comportement et le programmer ? C'est une question ouverte qui n'a pas encore été résolue, même si de nombreuses analyses faites depuis l'été 2015, à la suite de la lettre ouverte des chercheurs en intelligence artificielle sur les armes autonomes<sup>1</sup>, puis à l'automne 2015, après les débats sur les dilemmes moraux des voitures autonomes<sup>2</sup>, tendent à montrer la vacuité de beaucoup de ces inquiétudes.

Le troisième défi vient, non pas de la machine autonome elle-même, mais de son interaction avec l'homme et, plus généralement, de ce que l'on appelle les conflits d'autorité entre hommes et machines. En effet, bien souvent, on aspire à coopérer avec les robots, de sorte que les décisions soient, sinon collectives, tout au moins partagées. Cela veut dire que l'important tient à la coopération et à la répartition des tâches.

Or, l'expérience montre que les accidents avec les machines proviennent surtout d'une mauvaise coordination. C'est en particulier le cas dans le secteur aéronautique et nucléaire où l'étude rétrospective des catastrophes montre que, le plus souvent, l'action des hommes est en cause. Est-ce que cela signifie que les machines doivent toujours prendre le pas sur l'homme ? Si oui, comment accepter de donner systématiquement la prééminence à la machine sur l'homme ? Là encore, il n'existe pas de solution qui permette de trancher de façon définitive.

**Des personnalités, parmi lesquelles Stephen Hawking et Elon Musk, ont exprimé leur inquiétude face au danger potentiel de l'existence de machines devenues autonomes. Créatrice du logiciel AlphaGo, vainqueur au jeu de go face au champion mondial sud-coréen Lee Sedol en mars 2016, la société DeepMind (groupe Alphabet/ex-Google) s'interroge sur la capacité de l'intelligence artificielle à désobéir en empêchant l'interruption par l'homme d'un programme.**

**Qu'en pensez-vous ? Fantômes ou réalité ? La distinction entre une IA (intelligence artificielle) faible et une IA forte est-elle pertinente ?**

Beaucoup de personnalités publiques, jouissant d'une grande autorité dans le monde anglo-saxon, qu'il s'agisse de scientifiques de renom comme Stephen Hawking ou Ray Kurzweil, ou d'industriels richissimes comme Elon Musk ou Bill Gates, lancent de grandes déclarations publiques dans les médias pour manifester leurs inquiétudes face aux conséquences

ultimes des progrès époustouflants de l'intelligence artificielle. Leurs craintes reposent essentiellement sur deux arguments : le premier invoque la loi de Moore qui prédit l'amélioration exponentielle des performances des processeurs, ce qui conduira, selon eux, à fabriquer des machines dont les capacités excéderont bientôt celles des humains. Le second est relatif aux facultés d'apprentissage des machines qui outrepassent si fortement les nôtres qu'elles nous deviendront très bientôt incompréhensibles, au point que l'on ne saura plus les maîtriser.

Or, aucun de ces deux arguments ne résiste à une analyse sérieuse. En effet, la loi de Moore n'est qu'une loi d'observation ; elle n'a de fondement qu'empirique et ce n'est en aucun cas une loi de la nature. Elle s'est d'ailleurs vérifiée de façon plutôt erratique depuis une soixantaine d'années, avec des accélérations et des décélérations, puis de nouvelles accélérations suivies encore de décélérations, etc.

#### POUR ÉPOUSTOUFLANTES QU'ELLES SOIENT, LES CAPACITÉS D'APPRENTISSAGE DES MACHINES NE SONT PAS INFINIES

Aujourd'hui, il semble que les technologies du silicium avec lesquelles on réalise les circuits intégrés atteignent une limite physique de miniaturisation que l'on appelle le « mur du silicium ». Il se peut donc qu'on n'observe plus la loi de Moore, c'est-à-dire la progression exponentielle des performances, dans les prochaines années. Certains imaginent que de nouveaux paradigmes technologiques, comme celui du calcul quantique, permettront d'aller au-delà et de continuer à observer cette loi exponentielle d'accroissement des performances. C'est possible, du moins rien ne démontre le contraire, mais cela n'a rien de certain et, au vu des récents résultats obtenus par les physiciens, cela demeure fort peu probable.

Le deuxième argument tient aux capacités d'apprentissage des machines grâce auxquelles elles nous surpassent et nous surprennent. Ce fut, entre autres, le cas avec le programme AlphaGo de la société DeepMind qui l'a emporté sur l'un des meilleurs joueurs mondiaux du jeu de go en ayant recours à du *Deep Learning* et à de l'apprentissage par renforcement. Cela l'est aussi avec les véhicules autonomes ou avec les logiciels de reconnaissance vocale comme Siri, ou encore avec le système Watson qui l'a emporté sur les meilleurs humains au jeu de Jeopardy, même si, dans ce dernier cas, il n'y a pas vraiment d'apprentissage machine, mais utilisation de grandes quantités de connaissances.

À cet égard, notons qu'indépendamment des performances actuelles des ordinateurs dotés d'apprentissage, cela fait très longtemps que ceux-ci calculent plus rapidement et mieux que les hommes et qu'il y a bientôt vingt ans qu'un programme informatique l'a emporté sur le champion du

monde en titre au jeu d'échecs. En conséquence, la victoire d'un ordinateur sur l'un des meilleurs joueurs humains au jeu de go ne constitue pas une étape décisive dans l'histoire de l'humanité. Ajoutons à cela que, pour époustouflantes qu'elles soient, les capacités d'apprentissage des machines ne sont pas infinies. On sait qu'il existe des limites formelles qui bornent les performances des algorithmes d'apprentissage supervisé. De plus, l'apprentissage supervisé se restreint à un cadre conceptuel défini ; il n'invente pas de notions nouvelles. En d'autres termes, avec l'apprentissage supervisé, on se trouve dans l'incapacité à opérer des changements de paradigmes comme l'ont fait les grands découvreurs. Il faudrait certainement développer des algorithmes d'apprentissage non supervisé pour aller au-delà. Rien ne dit que les chercheurs n'y parviendront pas dans le futur, mais pour l'instant nous n'avons aucune certitude.

En somme, les inquiétudes manifestées par les personnalités influentes citées ici n'ont pas de fondements scientifiques. Dès lors, on peut se demander pourquoi elles se prononcent sur ces questions alors qu'elles n'ont pas nécessairement les compétences requises pour s'exprimer. Sans doute y a-t-il l'effet médiatique, car l'annonce de catastrophes rencontre toujours une grande audience. Mais, provenant de responsables d'industries engagés dans la production des technologies qui causent ces catastrophes, on peut s'interroger sur leurs motivations réelles.

#### IL CONVIENT DE RÉAFFIRMER QUE LE PROGRÈS TECHNOLOGIQUE N'A RIEN D'INÉLUCTABLE

À titre personnel, je crois qu'il y a là des enjeux politiques majeurs. En effet, ces craintes permettent de souligner à la fois le pouvoir des technologies, l'inéluclabilité du progrès et la sollicitude de grandes compagnies qui, comme Google, prétendent se soucier du devenir de l'humanité. Or, il convient de réaffirmer que le progrès technologique n'a rien d'inéluclable. D'ailleurs, dans l'histoire, on a connu de longues périodes de reflux technologique et des pertes de compétences majeures. De plus, le pouvoir des technologies n'est pas infini et, contrairement à ce que prétendent certains, on n'est pas encore en mesure de vaincre la mort...

Enfin, et c'est là certainement le point central, les grandes sociétés comme Google se soucient bien plus de leur avenir que de celui de l'espèce humaine...

Sources :

<sup>1</sup> <http://futureoflife.org/open-letter-autonomous-weapons/>

<sup>2</sup> <http://www.theatlantic.com/technology/archive/2015/10/trolley-problem-history-psychology-morality-driverless-cars/409732/>