

L'INTELLIGENCE ARTIFICIELLE, INTÉRÊTS, FANTASMES ET RÉALITÉS

Introduction : investissements et questionnements	page 2
L'Intelligence artificielle : une présentation historique	page 3
Le <i>Machine Learning</i>	page 4
Les réseaux de neurones artificiels	page 5
Le <i>Deep Learning</i> , apprentissage profond	page 6
Les algorithmes	page 7
Le "biais" d'un algorithme	page 8
Une intelligence pas si intelligente que ça...	page 10
Une illusion d'humanité	page 12
Les données numériques de la vie privée : indignation et vrais dangers	page 13
L'IA et la médecine	page 16
Sur l'intelligence forte	page 17
Les masques du transhumanisme	page 18
Ces robots qui fascinent les humains	page 22
Quel avenir pour le monde du travail ?	page 24
Quelques conclusions	page 27
Sources	page 28

Introduction : une montée en flèche des investissements et d'innombrables questions

Il y a dix ans à peine, l'intelligence artificielle était inconnue, juste une affaire de quelques spécialistes dans un domaine de la recherche qui semblait sans grand avenir. Aujourd'hui, on la découvre un peu partout, et on est régulièrement surpris de réaliser que nous l'utilisons, sans qu'on ne nous en ait jamais ni parlé, ni encore moins demandé notre autorisation pour devoir l'utiliser. On a donc une intelligence artificielle qui nous répond au téléphone, une autre qui peut nous aider à garer notre voiture, et demain -nous dit-on- à la conduire à notre place. On en a une qui peut écrire ce que nous disons, ou qui traduit en une fraction de seconde des pages entières d'internet dans notre langue, ou dans n'importe laquelle.

On ne nous a pas prévenus de la présence de l'intelligence artificielle. Mais au moins, la question que nous pouvons nous poser, c'est de savoir vers où cela nous mène-t-il ? Nous allons le voir au cours de cet exposé, on n'en sait en fait rien. Tout le monde n'a pas le même avis sur la question. Il y a la vision qu'en ont les principaux acteurs, il y a les craintes du public le plus large, et il y a encore celles de ceux qui étudient la question.

La seule chose qui est sûre et certaine pour l'heure c'est que l'argent afflue dans ce secteur. Plus de 13,5 milliards d'euros vont être investis en Chine pour le développement de l'intelligence artificielle pour les années 1919/1921. Et pour un peu plus tard, à l'horizon 2030, la Chine a déjà prévu un programme cette fois de 150 milliards de dollars, avec l'objectif de devenir leader mondial dans ce domaine. L'Europe estime qu'il faudrait investir 30 milliards en comptant le public et le privé. Elle compte déjà 2260 start-up et 380 laboratoires qui travaillent dans l'intelligence artificielle.

En France, comme dans bien d'autres domaines, on est très discret sur la puissance réelle du pays. Macron a annoncé un plan de 1,5 milliard sur quatre ans. Mais il faut savoir que, d'ores et déjà, 268 équipes de chercheurs travaillent sur l'intelligence artificielle, ce qui totalise 5300 chercheurs. Dans les grandes écoles et les universités, il y a actuellement 138 cours différents liés à l'intelligence artificielle et 18 diplômes de masters spécialisés, qui vont jusqu'au droit de l'intelligence artificielle ou à l'éthique. Au niveau des entreprises, on compte 300 start-up spécialisées en IA (intelligence artificielle) et 80 entreprises de plus grande taille, entre 50 et 5000 salariés. Ce qui place la France à la quatrième place mondiale dans ce domaine, derrière la Chine, les USA et le Royaume-Uni.

Mais les changements de comportements induits par l'IA posent de nombreuses questions. Ainsi, comme le souligne l'écrivain israélien Yuval Noah Harari, *“en l'espace de deux décennies à peine, des milliards de personnes en sont venues à confier à l'algorithme de recherche de Google l'une des tâches les plus importantes qui soient : trouver des informations pertinentes et fiables (...). Nous comptons de plus en plus sur Google pour ses réponses, notre capacité à trouver des informations de manière autonome diminuée”*.

Comment tout cela va-t-il impacter la vie humaine ? la conduite automobile, le travail, le chômage ? les humains peuvent-ils être dépassés par l'intelligence artificielle ? Et cette

intelligence artificielle, n'y a-t-il pas derrière elle des humains qui la contrôlent, qui peuvent donc nous manipuler ?

L'Intelligence artificielle : une présentation historique

L'intelligence artificielle est née d'une autre direction de recherche, que l'on peut dater de l'après-guerre, dans un domaine appelé alors la cybernétique. Aujourd'hui, le qualificatif cyber fait allusion à ce qui est virtuel, numérique, à internet. Entre 1946 et 1953, en effet, des chercheurs aussi bien en mathématiques, des ingénieurs, des physiologistes, des anthropologues, se réunissent pour travailler autour d'un projet commun, qu'ils dénomment cybernétique : il s'agit de comprendre et de développer des systèmes automatiques qui utilisent les transmissions d'information, que ce soit pour des machines, ou chez les organismes vivants.

L'"intelligence artificielle" est une fille de ces premières recherches. Déjà en 1952, l'américain Arthur Samuel avait conçu un programme sur l'ordinateur IBM 701, qui permettait à la machine de jouer au jeu de dames avec un joueur de bon niveau. Cette réussite semblait déjà extraordinaire pour l'époque. En 1956, une conférence se tient au Dartmouth college, aux États-Unis, qui réunit des chercheurs en cybernétique, des chercheurs en biologie des neurones, des spécialistes du traitement de l'information. Leur idée est que chaque aspect de l'intelligence humaine, et l'apprentissage notamment, peut être décrit de manière assez fine, de sorte que l'on puisse le reproduire dans une machine. Il s'agit donc de chercher à créer une "intelligence artificielle". Le terme apparaît alors, aujourd'hui on utilise couramment l'abréviation IA.

De cette année 1956 jusqu'aux années 1970, c'est l'enthousiasme pour l'intelligence artificielle. C'est à celui qui fera la déclaration la plus optimiste. *“D'ici dix ans, un ordinateur sera le champion du monde des échecs”*, promettent H. Simon et Allen Newell en 1958. *“Des machines seront capables d'ici vingt ans, de faire tout travail que l'homme peut faire”* : c'est encore H. Simon en 1965. *“Dans trois à huit ans, nous aurons une machine avec l'intelligence générale d'un être humain”* : c'est Marvin Minsky en 1970. Marvin Minsky avait en tout cas réussi à construire la première machine à réseau neuronal, sur laquelle nous reviendrons dans un instant. Dans ces conditions, très optimistes, l'argent afflue vers ces chercheurs, que ce soit de la part des gouvernements ou des militaires.

Mais, fin des années 1970, on finit par voir la réalité en face : on n'a pas beaucoup avancé, pour ne pas dire pas du tout. Les critiques se font virulentes contre ces recherches apparemment inutiles, et l'argent va alors cesser de couler à flots. C'est la crise pour l'IA.

Il faudra attendre plus d'une vingtaine d'années, au milieu des années 1990, pour qu'un renouveau de l'IA se manifeste. Cette fois, on est moins ambitieux, il ne s'agit pas de copier un être humain tout entier, mais de travailler sur des domaines particuliers, comme les robots, notamment dans l'industrie. Et ça fonctionne. En 1997, le champion du monde des échecs Gary Kasparov est battu par le supercalculateur IBM Deep Blue.

C'est que presque tout a changé, par rapport aux années 1970. On dispose de vitesses de calcul qui ont été multipliées par un million. On dispose ensuite de masses de données considérables, qu'on n'avait pas du tout dans les années 1970. Et surtout, on ne cherche plus du tout à reproduire, dans le programme donné à la machine, tout ce que fait un humain.

L'idée même d'intelligence artificielle, avec une sorte d'analogie avec l'intelligence humaine, va rester absente dans le domaine de la science sérieuse, jusqu'au milieu des années 2000.

Et puis, autour de 2005/2010, soudain, une véritable révolution va se produire, et cette fois se développe à toute vitesse, s'étend à tous les domaines, permettant de mettre au point des systèmes de caméras qui reconnaissent les visages, des téléphones portables à qui vous pouvez dicter un texte, et qui le transcrivent par écrit, ou qui le traduisent instantanément en une multitude de langues, des systèmes capables d'analyser des images médicales pour détecter la présence de tumeurs cancéreuses avec un meilleur taux de réussite qu'un médecin humain. Cette prouesse technique est due à la récupération d'une vieille idée des années 1970, l'idée des neurones artificiels, inaugurée par Marvin Minsky.

Deux grandes étapes vont marquer l'histoire récente de l'intelligence artificielle : le *Machine Learning* avant 2005/2010 et le *Deep Learning* ensuite.

Le *Machine Learning*

Imaginons, par exemple, que l'on veuille développer un programme qui soit capable de reconnaître l'image d'un avion. Si je dis à un humain d'imaginer un avion dans sa tête, et que je lui en présente ensuite une image, il y aura évidemment des différences, mais pour lui, dans les deux cas, c'est un avion. L'idée générale et le cas particulier rencontré sont différents. L'humain cherchera par exemple si l'image a bien deux ailes, une cabine de pilotage, etc.

Dans le *Machine Learning*, on va apprendre à la machine ce genre de choses. Et on va faire en sorte qu'on puisse agir sur la machine en différents points de son programme, que j'appellerai pour simplifier ici des "boutons". Il y aura un bouton pour reconnaître les deux ailes, un autre pour les roues, etc. On a ainsi fabriqué, sans trop savoir ce qu'il vaut, un début de programme de reconnaissance d'avion. Pour vraiment le mettre au point, on va présenter à la machine des milliers d'images et lui demander s'il s'agit ou non d'un avion. Et chaque fois qu'il y a une erreur, on va agir sur un des "boutons" pour en modifier la position, pour modifier le programme.

Dans ce travail, ce qui sera essentiel, ce sera de distinguer, dans ce que nous voyons, des caractéristiques discriminantes, des points sur lesquels on fera la différence entre un avion, un oiseau (qui a aussi des ailes), un vélo (qui a aussi deux roues), une auto (qui a aussi un poste de pilotage), etc. On peut ainsi trouver peut-être dix ou quinze points qui permettent de juger et de décider si on a affaire à un avion ou pas. En fait, on a résumé l'image à ces formes essentielles.

Ainsi, progressivement, on va améliorer le programme, et on aura un taux de réussite de plus en plus élevé. Mais cette méthode a un gros inconvénient, c'est qu'elle demande une intervention humaine très lourde pour apprendre, en fait, à la machine. On appelle aussi cette IA une IA "supervisée".

Comment, dans la pratique, obtient-on cette situation dans une machine où des positions différentes peuvent être modifiées ? On fabrique ce qu'on appelle des réseaux de "neurones artificiels".

Dans le cerveau humain, chaque neurone peut recevoir quelque chose comme 100 000 signaux électriques venant d'autres neurones. La communication d'un neurone à un autre se fait à un endroit particulier, qu'on appelle la synapse. En fonction de ce qu'il a reçu, le neurone va ensuite soit ralentir, stopper le message qu'il a reçu, soit au contraire l'exciter, le propager. Plus un neurone est sollicité, plus il "apprend", plus il aura de "force synaptique", plus il aura de chances d'envoyer des signaux à ses voisins quand il sera stimulé par une impulsion électrique. C'est donc le nombre de fois que le neurone a reçu un signal qui compte dans son fonctionnement.

On va donc essayer de reproduire ces principes dans une machine, même si c'est sous une forme extrêmement simplifiée. Déjà en 1943, les mathématiciens Mc Culloch et Pitts avaient réalisé le premier modèle simple d'un neurone : un automate qui représente le noyau de la cellule, l'axone (qui relie le centre du neurone à son lieu de connexion) et les dendrites (les connexions avec les autres neurones). Et à partir d'un certain seuil d'activation, le neurone envoie de l'information aux neurones qui lui sont connectés.

On a donc repris l'idée des années 1970 de Marvin Minsky, et on utilise le principe avec lequel le cortex visuel - la partie du cerveau humain qui traite les informations-, analyse les signaux venant des yeux. Et là, les neurosciences nous ont appris une chose, c'est que les neurones sont organisés en couches.

Les signaux lumineux qui atteignent les cellules de la vision dans l'œil humain sont transformés en impulsions qui arrivent d'abord dans une première couche de neurones. De cette couche, sort un ensemble d'informations qui vont ensuite vers une seconde couche, et ainsi de suite. Chaque couche a des neurones qui ont une histoire différente, un passé différent, et qui vont réagir selon ce passé, c'est-à-dire selon la quantité des anciens signaux qu'ils ont reçus ou pas. Les premières couches se concentrent sur les détails, aux échelles les plus petites. Les suivantes, différentes donc des premières, vont analyser et suivre les informations reçues des couches précédentes à une échelle plus large. Finalement, plus on avance de couche en couche, plus les couches donnent une représentation abstraite, générale, de ce que l'on a vu avec les yeux. Jusqu'à la couche finale qui donne au cerveau une image finale de la situation.

En 1986, on est parvenu à construire un premier réseau de neurones artificiels formé sur plusieurs couches. On a aussi mis au point un programme - un algorithme - appelé de "rétro-propagation du gradient" - qui permet d'entraîner ces réseaux de neurones. Mais, jusqu'en 2005, les tentatives faites de construire une telle intelligence artificielle en utilisant plus de deux ou trois couches ont échoué. En fait, on faisait faire le même travail par toutes les couches. Mais à partir de 2005, on a réussi à faire travailler différemment les couches les plus profondes. Et cela a commencé à marcher.

Deux autres facteurs ont joué dans les progrès qui sont alors survenus. D'abord, la puissance de calcul des ordinateurs grâce aux processeurs : et l'on est passé, en moins de 30 ans, de 20 millions d'opérations (addition, soustraction, etc.) par seconde à des millions de

millions d'opérations par seconde. Le second facteur est qu'on a maintenant d'énormes quantités de données, des millions ou plus par exemple d'images de chat, pour enseigner à la machine à reconnaître un chat. Or, plus on en a de ces données, mieux on apprend à la machine, et moins elle se trompera.

Le *Deep Learning*, ou apprentissage profond

Mais, si vous avez bien suivi notre exposé, il reste le problème de la supervision : un humain, ou des humains doivent encore dire à la machine qui apprend : ça c'est un chat, ça, ce n'en est pas un ; ou ça c'est le son A, ça ce n'est pas le son A. Et faire cela pour des millions d'images, c'est fastidieux. C'est une grosse limite.

Le problème est aujourd'hui réglé. La solution ? C'est la machine qui apprend toute seule. Les quantités gigantesques de données dont on dispose et les puissances de calcul phénoménales dont on dispose font que l'on peut maintenant laisser la machine apprendre seule : elle s'entraîne seule à reconnaître les caractéristiques spécifiques des objets. Elle apprend donc seule dans une première étape. Il lui suffit de chercher sur Google tous les endroits où il y a à la fois la mention du mot chat et une image correspondante. Après quoi, elle est prête pour servir et faire le travail. Ainsi, en 2012, le projet *Deep Learning* de Google, nommé Google Brain, a été capable, après trois jours de travail, sans intervention humaine, en utilisant dix millions de captures d'écran faites de manière aléatoire sur You Tube, de découvrir par lui-même les concepts essentiels permettant de distinguer un chat ou une silhouette humaine.

Mieux, alors que dans l'étape précédente, que l'on a vu avec l'exemple de l'avion, on se contentait de repérer quelques caractéristiques dans l'image, on peut maintenant donner à l'ordinateur qui travaille l'image complète à analyser. Une image de photo ou sur machine est formée d'un très grand nombre de points, les pixels. Sur un rectangle de 960 pixels de large et 674 pixels de haut, ça nous donne 647 000 briques d'information à traiter. Impossible à faire il y a quelques années. Maintenant, on le peut.

La consécration du Deep Learning ("apprentissage profond", profond parce qu'on va plus loin dans la profondeur des couches de neurones) a eu lieu avec le match, en octobre 2015, entre le Français (d'origine chinoise) champion d'Europe du jeu de go, Fan Hui, et le logiciel AlphaGo de Google. Fan Hui s'était déjà affronté à des logiciels, toujours vainqueurs, rigolant même de certaines erreurs grossières faites par les machines. Et même DeepBlue d'IBM, qui avait gagné aux échecs, ne faisait pas le poids au jeu de go. Han Fui était donc tranquille. Il a été écrasé par DeepBlue, par 5 à 0.

Un pas a donc bien été franchi. Comment ? On a fait analyser par AlphaGo 30 millions de parties de go jouées par des professionnels. On a défini les critères à sélectionner pour les suivre en priorité, et trouver les coups les plus efficaces. Après quoi, on a laissé le programme jouer contre lui-même, sur 50 ordinateurs différents. Ce qui améliorait à chaque fois sa stratégie de jeu. AlphaGo a ensuite battu le champion du monde, Lee Sedol, en 2016.

Et, bien évidemment, dans le système capitaliste où chacun est concurrent des autres, dès les premiers résultats obtenus, une pluie d'argent, par millions, s'est abattue sur tout ce qui a un rapport avec l'IA : il s'agit de se placer le plus en avance possible sur les concurrents.

Mais attention, que la machine apprenne seule, cela ne veut pas du tout dire que les humains n'existent plus. C'est un travail humain qui construit le programme que suit la machine pour faire les recherches, trier, chercher, déterminer ce qu'elle retiendra ou pas. Ce programme, traduit dans un langage que suit l'ordinateur, a pour nom en informatique un "algorithme".

L'algorithme est une suite d'instructions très simples, un peu comme une recette de cuisine, lisibles et effectuables par l'ordinateur. Dans l'exemple qui suit, l'algorithme calcule, puis affiche, l'âge qu'aura une personne en 2100 :

```
variables
annee, age, age_en_2100 : entiers
début algorithme
afficher "Entrer l'année actuelle"
lire année
afficher "Entrer votre âge"
lire âge
âge_en_2100 prend la valeur âge + 2100 - année
afficher "En 2100, vous aurez ", âge_en_2100, " ans."
fin algorithme
```

On donne d'abord à la machine la liste des variables qui vont être utilisées dans cet algorithme. On commence par lui dire d'afficher à l'écran la phrase "Entrer l'année actuelle" ; puis on lui demande de lire ce que la personne a marqué ; ensuite on lui demande d'écrire la phrase : "Entrer votre âge," et ainsi de suite.

A la fin, si, par exemple, l'utilisateur a entré "2014" comme année actuelle et "16" comme âge, l'algorithme affichera : *En 2100, vous aurez 102 ans.*

Cet exemple est très simple, mais l'idée est là : il s'agit toujours de décomposer les ordres de manière aussi simple.

S'il s'agit de calculer un âge, rien ne peut nous influencer dans notre manière de voir, et l'algorithme est donc parfaitement neutre. Mais pour bien des questions posées à l'intelligence artificielle, consciemment ou pas, le point de vue de celui qui construit le programme et l'algorithme peut jouer.

Ainsi, en 2016, Microsoft (Word, Excel, PowerPoint, etc.) a construit un agent conversationnel, nommé Tay. Tay devait simuler être une jeune adolescente qui discute naïvement sur Twitter. Et chacun pouvait donc lui enseigner ce qu'il voulait, puis discuter avec elle. Tay avait commencé son expérience en disant des choses du style "*les humains sont super cool*". Mais des internautes ont décidé de lui apprendre cette chose qu'on trouve facilement sur internet, la violence verbale. En vingt-quatre heures, Tay a alors posté des messages comme : "*Je hais les féministes, elles devraient toutes brûler en enfer*", ou bien "*Hitler a fait ce qu'il fallait*". Microsoft a dû arrêter son expérience.

Un peu comme dans une recette de cuisine, un algorithme va d'abord indiquer à la machine les ingrédients à utiliser, après quoi, on aligne une liste d'instructions qui peuvent

faire une demi page, deux pages, ou quatre mille pages. En cuisine, ce sera : chauffer l'eau, cuire les pâtes, découper les tomates ; en informatique c'est : compter de tant à tant, enlever ceci, trier cela selon tel critère. Les ingrédients, cela peut être des nombres, mais aussi des images, des sons, etc.

Si on veut construire un algorithme qui rend plus nette une image donnée floue en noir et blanc, on écrira des instructions pour dire à la machine : lire le premier pixel, s'il est gris clair, le rendre plus clair, s'il est gris foncé le rendre un peu plus foncé ; passer au second pixel, recommencer. Mais on peut compliquer un peu : si le pixel est sur un bord, rendre plus clair encore, si le pixel est au centre, ne rien changer, garder la finesse. On peut encore plus compliquer : si un pixel moyennement gris se trouve entouré de nombreux pixels noirs, le rendre noir, lui aussi ; s'il est moyennement gris et entouré de pixels nombreux blancs, le rendre blanc ; etc.

Mais on voit ici du coup qu'un algorithme destiné à une fonction ne pourra pratiquement rien faire d'autre que cette fonction. Une machine à qui on aura appris à conduire une voiture toute seule ne comprendra rien si elle voit une cuisine autour d'elle. Elle ne pourra même pas aider le conducteur à sortir de sa voiture.

En fait, à peu près tout ce qui est fait actuellement en intelligence artificielle a cette limite : l'engin n'est pas autonome, et il est incapable de s'adapter à une autre situation que celle pour laquelle on l'a conçu.

Mais que cela ne se fasse pas, n'existe pas, n'a pas empêché des chercheurs ou des vantards d'extrapoler, et d'affirmer qu'on va tout droit dans cette direction, qu'on y arrivera dans dix, 20, 30 ans. Ils appellent cela l'intelligence artificielle "forte", celle qui sera capable de s'adapter, comme le fait en permanence un être humain, ou même un animal.

En attendant, aujourd'hui, il n'existe que de l'intelligence artificielle "faible". Et cela nous pose déjà des problèmes.

Le "biais" d'un algorithme

Nous autres, les humains, et sans doute aussi bon nombre des autres animaux, nous pouvons aussi avoir un fonctionnement qui ressemble à un algorithme, même si nous ne nous en rendons pas compte. Par exemple, si nous nous posons la question d'aller au cinéma voir un film, après en avoir lu une critique dans le journal. Nous hésitons, nous ne savons pas trop. Alors, nous nous posons une série de petites questions simples : est-ce que je connais quelqu'un qui a vu ce film ? Est-ce que je connais son avis ? Est-ce que j'ai aimé les films que j'ai vus de ce réalisateur ? ou avec cet acteur ? Est-ce que j'ai envie de voir un policier, ou tel genre de film ? Est-ce que je n'ai pas plutôt envie d'aller me promener, pour une fois qu'il fait beau ? et ainsi de suite, on passe par une série de questions réponses : chacune est simple, c'est la réponse finale qui n'est pas évidente. C'est un peu un algorithme.

Comment est prise la décision ? En mettant des valeurs différentes pour chacune des questions. On va par exemple juger que l'avis sur le réalisateur compte peu, disons si on veut mettre un chiffre 1 sur 10, que l'acteur compte plus, mettons 3 ; que l'avis d'un copain, compte plus encore, 4, que le beau temps vaut pour 2. Bien sûr on ne pense pas à des chiffres de cette manière, mais on pèse et on soupèse les choses un peu quand même. Le résultat va dépendre

au final de ce comptage, même si on le fait au jugé. Et peut-être que notre cerveau, lui, fait vraiment une forme de comptage, pas si au jugé que cela.

Nous en arrivons maintenant au type de problème que peut poser la construction d'un algorithme. Imaginons, et cela existe, un algorithme destiné à décider du recrutement d'un personnel qualifié. Pour le construire, on va d'abord lister toute une série de facteurs, dont on pense qu'il faudra en tenir compte : l'âge du candidat, son expérience professionnelle, son sexe, son adresse, ses diplômes, et toute une série de qualités que l'on peut lui trouver ou pas, à partir de son CV ou lors d'un entretien de recrutement. Mais ensuite, il faut décider, comme nous pour le cinéma et l'avis du copain, du poids, de l'importance que l'on va donner à chaque facteur.

La machine, elle, ne sait faire que cela, compter, et il lui faut des chiffres. Question : est-ce qu'on donnera le même poids selon qu'il s'agit d'un homme ou d'une femme ? Certains féministes vont hurler que oui. Mais si c'est une sage-femme qu'on veut recruter ? Ou si au contraire, c'est un boucher, qui devra porter des carcasses de 50 kilos ? Si c'est vous le programmeur, et si vous n'avez pas envie de casser le dos d'une dame, vous allez peut-être décider de mettre 1 point si le candidat est une femme, contre 2 pour un homme... Et ainsi de suite pour chacun des critères.

On voit que, selon la manière dont vous raisonnez, vous allez fabriquer un algorithme qui peut donner des réponses différentes, des réponses liées à votre manière de voir. Donc que, derrière les apparences strictement mathématiques, logiques, apparemment neutres, qu'a l'algorithme une fois écrit, il y a bel et bien une vision sociale qui y est imprimée.

Comment sont réalisés les algorithmes actuels ? Avec quelle vision sociale ? Le plus probable, c'est qu'ils veulent surtout éviter les problèmes, pour continuer à être acceptés et vendus. On peut donc penser qu'ils sont plus ou moins le reflet de la société telle qu'elle fonctionnait déjà avant l'algorithme. Sauf que cette société pose de bien nombreux problèmes. Que font les algorithmes ? On nous dira sans doute qu'ils sont neutres, transparents. Non, ils sont opaques, et reproduisent ce qui existe déjà, et qui est déjà souvent injuste.

Ce problème est tellement présent en intelligence artificielle, qu'on lui a donné un nom : cette déformation induite par le programmeur quand il fait passer sa manière de voir dans l'algorithme, on appelle cela un "biais" : le programme est biaisé, tordu, dans son sens.

Et des biais, il peut y en avoir bien avant la construction de l'algorithme. Pour reconnaître des visages de personnalités, par exemple, vous savez maintenant qu'il faut montrer à la machine des milliers et des milliers d'images de chacune d'entre elles, en lui disant son nom. Mais les choix des ingénieurs dès cette étape peuvent avoir des conséquences.

On s'est ainsi aperçu que les personnalités féminines sont moins bien et moins facilement reconnues que les personnalités masculines. Une étude du MIT de février 2018 a montré que les logiciels d'IBM, de Microsoft et de Facebook, lorsqu'on leur demande de reconnaître 1270 portraits officiels, ont des résultats nettement meilleurs pour les hommes (Facebook en a reconnu 99,3%) que pour les femmes (seulement 78,7%).

Sexiste, l'algorithme ? Les ingénieurs qui ont préparé le travail, oui ! Mais ils peuvent être aussi racistes : les logiciels de reconnaissance faciale ont plus de mal à distinguer le visage d'un noir que celui d'un blanc. Autre exemple, dans les logiciels de traduction : le mot

"nurse" (qui est neutre en anglais) est traduit par infirmière (au féminin, alors que le neutre "infirmier" existe). Tandis que le mot "doctor" (également neutre en anglais) est traduit par docteur (bien au masculin, lui).

Ce genre de biais, on peut penser qu'ils sont involontaires, qu'ils sont simplement le reflet de la manière de voir "normale" de ceux qui ont travaillé. Ils n'en sont pas moins problématiques. Un autre exemple nous est donné par Nicolas Spatola, qui travaille sur les interactions homme-machine à l'université de Clermont : *“Des offres d'emploi de Google gérées par l'intelligence artificielle proposaient pour un même emploi, des revenus différents selon le genre du candidat. Ce biais, nous dit-il, provenait de la présence d'inégalités constatées dans la société lors de l'apprentissage de l'intelligence artificielle”*.

Une intelligence pas si intelligente que ça...

Avant d'aller plus loin, de regarder les applications pratiques de l'IA, et de réfléchir aux changements que cela peut impliquer dans la vie sociale en général, nous allons tout de suite mettre une limite à l'idée d'"intelligence artificielle".

Même si la machine peut maintenant apprendre sans l'utilisation d'un homme qui lui dit *“ceci est un chat”*, *“ceci n'est pas un chat”*, elle a besoin, on l'a dit, d'avoir été "voir" des quantités phénoménales de chats. Or, le moindre enfant humain, n'a besoin de voir qu'un chat ou deux, pour, définitivement, savoir reconnaître un chat en toutes circonstances, et quasiment jamais se tromper. La machine, elle, qui a appris d'une toute autre manière, peut avoir des ratés. Certaines images de chats ne sont pas reconnues comme un chat, et inversement, il n'est pas impossible qu'une image floue style tâche d'encre soit pour elle un chat, alors qu'elle ne tromperait aucun être humain.

C'est que la machine apprend, ici, d'une manière qu'on appelle "ascendante" : à partir d'une profusion d'exemples, elle "remonte" vers une idée générale.

Mais les humains, eux, combinent cette manière d'apprendre avec une autre, bien plus économique, une approche "descendante". À l'école, par exemple, on nous explique que le A majuscule est formé par un angle en haut, et une barre au milieu. Du coup, vous reconnaîtrez un A majuscule même s'il est déformé vers la droite ou la gauche, ou légèrement courbe, et ce dans de très nombreuses polices de caractères. Pas besoin de montrer à l'élève deux millions de A, ni même 200. Son cerveau est déjà tout à fait capable de fonctionner sur la base d'abstractions. Et il sait aussi les relier entre elles.

La méthode descendante, c'était un peu celle des premières tentatives d'IA, dans les années 1950/1960. Mais elles avaient de grosses limites : dans la manière de discriminer, par exemple : il faut reconnaître deux ailes de l'avion, mais il se peut que dans certaines positions on n'en voit qu'une.

Aujourd'hui, des chercheurs travaillent sur la manière dont les enfants apprennent, dans l'espoir de pouvoir imaginer, plus tard, de nouvelles méthodes pour élaborer l'IA. On essaye d'étudier comment les enfants apprennent les relations de cause à effet, ou encore quand et comment ils changent de point de vue sur une conviction erronée. En 2015, à l'université de New York, une étude a expliqué avoir réussi à apprendre à la machine à reconnaître des caractères manuscrits inhabituels, chose très simple pour les humains (même

si on ne connaît rien à la langue en question, on reconnaît l'arabe ou le chinois), mais extrêmement difficile pour les ordinateurs.

La méthode descendante n'a pas besoin de millions d'exemples, mais elle ne peut bien marcher que si on a vraiment trouvé les bonnes caractéristiques, elle demande donc un bien plus gros travail avant de démarrer.

Les enfants, eux, combinent les avantages des deux méthodes. Ils peuvent savoir ce qu'est un chat avec un seul chat, surtout s'il est vivant. Ils peuvent aussi faire des liens entre différentes choses qu'ils ont apprises pour en déduire, seuls, une connaissance nouvelle.

D'ailleurs, après une première période d'exaltation sur les réussites de l'apprentissage profond, certains chercheurs se sont amusés à chercher les défauts du système et ils en ont trouvé. Ainsi, ces machines qui surpassent de très loin l'être humain en capacité de calcul, en vitesse, en quantité de données qu'ils sont capables d'ingurgiter, en fiabilité de leur mémoire, etc. ne sont en fait pas si intelligentes que leur appellation le dit.

Un des pionniers de l'apprentissage profond, qui a dirigé la recherche en IA chez Facebook, Yann LeCun, est clair : *“C'est un abus de langage que de parler de neurones ! De la même façon qu'on parle d'aile pour un avion, mais aussi pour un oiseau, le neurone artificiel est un modèle extrêmement simplifié de la réalité biologique”*. Et il conclut : *“Les intelligences artificielles les plus abouties ont aujourd'hui moins de sens commun qu'un rat !”*

Effectivement ! Dès 2014, des équipes de chercheurs ont commencé à mettre au point des méthodes qui trompent toutes les intelligences artificielles actuelles, dans la reconnaissance visuelle. On pensait que si un ordinateur avait appris à reconnaître un cheval, il le reconnaîtrait même si on enlevait quelques pixels, sur les milliers que l'image comporte. Eh bien, pas du tout ! Enlevez quelques pixels à certains endroits, et l'IA va voir une vache au lieu du cheval. On pensait qu'une image proche d'une autre image correctement traitée sera aussi correctement traitée, mais ce n'est pas le cas.

Dans certains cas, un seul pixel ajouté peut même suffire à créer ce qu'on appelle une image piège pour la machine. Avec la même méthode, on peut partir d'une parole qui dit "Yes", oui en anglais, pour y introduire, de manière insensible à l'oreille humaine, une modification, qui fera entendre "No", non, par l'intelligence artificielle. À partir de là, certains envisagent de faire de ces méthodes des armes contre un ennemi dans une cyber guerre.

Mais, comme pour tout, les humains trouvent une parade à la parade. Par exemple, si on compresse une image, on a de bonnes chances d'enlever les pixels parasites et piégeux. Car on se ramène à l'essentiel de l'image.

En tout cas, nous voyons dès à présent que l'apprentissage dit profond est peut-être profond par ses neurones artificiels, mais pas du tout par la nature de ce qui est appris ou maîtrisé. Il peut très bien dire n'importe quoi.

Le *Deep Learning*, l'apprentissage profond, - le top du top actuellement - ne sait pas enchaîner les raisonnements ; il est complètement incapable par exemple de traiter des situations complexes, comme le fonctionnement d'un moteur de voiture ; il est même incapable de suivre un raisonnement enchaîné du genre : *“Lucie est la mère de Jacques, et la fille de Jean, donc Jacques est le petit-fils de Jean”*. L'apprentissage profond est incapable de

distinguer une corrélation (deux évènements se produisent en même temps, mais c'est peut-être juste une coïncidence, ou les deux évènements sont dus à un troisième que je ne connais pas) d'une cause (un évènement en produit un autre et en est le responsable). Enfin, et c'est un gros problème que l'on reverra plus loin, l'apprentissage profond est opaque : quand la machine vous donne un résultat, on est incapable de savoir pourquoi elle a choisi cette réponse, plutôt qu'une autre.

Une illusion d'humanité

En 2018, Google a sorti Google Duplex, une interface vocale, en clair une intelligence artificielle capable de discuter par téléphone, avec un salon de coiffure ou un restaurant, pour vous prendre un rendez-vous. Des voix artificielles, ce n'est pas vraiment nouveau, nous y sommes habitués, que ce soit pour nous annoncer la prochaine station dans le métro, ou quand on tombe sur une plate-forme téléphonique : c'est mécanique, froid, monotone. Mais là, avec Google Duplex, les personnes qui assistaient à la démonstration étaient médusées : le ton, les inflexions de la voix, et même des hésitations comme quand on fait "mmm euh", tout y était, comme une vraie voix humaine ! Et à l'autre bout, la personne pensait avoir un vrai client au téléphone.

Google a-t-il percé le mystère de la vie qu'il y a dans notre voix ? Et si c'est le cas, Google n'est-il pas, petit à petit, morceau par morceau, en train de préparer un futur être humain complet, avec son intelligence artificielle ? Nous allons voir que non seulement Google en est très loin, même si toute sa manière de faire voudrait nous le faire croire, mais qu'il n'a même pas commencé à comprendre un dixième de ce qui fait la voix humaine.

Ce que nous mettons dans notre voix est considérable. Jean-Luc Schwartz, directeur de recherche au CNRS Grenoble, énumère : *“Pour parler, les humains produisent du son par les cordes vocales, activées par le souffle de leurs poumons. Le son est modulé par le conduit vocal qui bouge rapidement sous l'action de la mâchoire, de la langue, des lèvres. Les modulations sont contrôlées par notre cerveau et nos connaissances sur le langage, pour produire, justement, des phonèmes (des consonnes et des voyelles), des mots et des phrases. Nous contrôlons aussi notre prosodie, le “chant de notre voix” (avec des montées et des descentes, des accélérations et des pauses) et puis nous y glissons des tics de langage, des petites particularités qui nous caractérisent - et que reproduisent à merveille les imitateurs.”*

“Le résultat de tout ce processus fait le grain de notre voix, son identité, ce son qui reflète à la fois notre manière de respirer, les fines particularités de nos cordes vocales, la forme précise de notre bouche et de notre palais, l'anatomie de notre mâchoire et de notre langue, la plasticité de notre peau, mais aussi le long cheminement de notre apprentissage de la parole depuis notre naissance et à travers notre enfance, notre histoire de chanteur ou de fumeur, et cette manière si particulière que nous avons de réfléchir, d'acquiescer, de grommeler...”

Que fait Google Duplex de tout cela ? Rien. Ou plutôt, une seule chose : pour que son langage soit compréhensible, il reproduit ce qu'on a compris sur le fonctionnement du langage, sur sa structure en phonèmes et en mots, sur les principes qui font que telle série de sons prend tel sens. Sur tout cet aspect, l'intelligence artificielle travaille, et produit un résultat qui permet que l'on comprenne ce qui est dit. Mais c'est tout. Alors d'où lui viennent ces hésitations, ces intonations ?

Là, Google utilise la base de l'intelligence artificielle, que nous avons décrite plus haut. On a stocké des milliers d'exemples de mots, de phrases. Et c'est dans ce stock que l'IA va redécouper des petits bouts pour produire sa voix soi-disant humaine. Google Duplex imite, mais ne ressent rien. Il apprend sans comprendre. Ceux qui l'ont programmé ont juste trouvé un moyen de faire semblant. Et si de nouvelles situations existent dans la vie, il faudra les lui apprendre, lui ingurgiter de nouvelles phrases, de nouvelles expressions. Mais l'illusion nous trouble, comme tout ce qui se rapproche de l'humain nous trouble. Google le sait, et en joue.

Autre exemple d'illusion, dans le journal Le Monde. En mars 2015, à l'occasion des élections cantonales, le journal a commencé à utiliser une intelligence artificielle, qu'on peut donc appeler des robots-rédacteurs, pour rendre compte des résultats : Data2Content a ainsi rédigé 36 000 petits articles, qui en réalité étaient pondus par l'IA pour donner un compte rendu en texte à partir des résultats chiffrés officiels.

Les données numériques de notre vie privée : l'indignation et les vrais dangers

Nous avons vu que l'un des facteurs qui a permis l'explosion de l'intelligence artificielle, c'est l'utilisation massive de données par la machine pour qu'elle puisse apprendre et bien apprendre. Comme on n'a plus trop de problèmes de capacités de travail, grâce aux performances des micro processeurs, on a maintenant plutôt des difficultés pour collecter des quantités énormes de données.

Nous avons tous entendu et compris que des données qui nous concernent sont capturées, récupérées, et éventuellement stockées. Ce qui choque, le plus souvent, c'est qu'on a le sentiment que nous sommes suivis, fliqués. On ouvre une page sur internet, et voilà que surgit une publicité de chez Darty pour le produit sur lequel on a fait une recherche la veille sur internet.

On a le sentiment d'une intrusion dans notre vie privée. En réalité, il n'y a intrusion de personne. Personne ne s'intéresse à nous, en tant qu'individu dont on voudrait connaître individuellement les habitudes, les pratiques, les pensées. Pour l'instant, c'est une machine de chez Darty qui, de manière automatique, en faisant ce que son algorithme lui dit de faire, a noté que vous aviez fait une recherche sur ce produit, et vous le propose le lendemain. Son algorithme lui dit de le faire, mettons trois fois par jour, pendant dix jours. Et elle le fait. La machine qui fait cela ne sait pas qui vous êtes, elle ne fait pas le lien avec d'autres recherches, elle ne connaît pas tout ce que vous avez été faire sur tout internet. Et il est plus que probable que personne, aucun être humain ne va s'intéresser à vous, individuellement, pour savoir tout ce que vous avez déjà pu acheter, ou consulter, sur le site de Darty.

Il n'y a donc pas intrusion dans notre vie privée, en réalité. Mais il y a collectage de données, en vue de faire du commerce. C'est un collectage global, collectif, qui amasse les recherches de tous ceux qui vont sur le site, pour leur proposer des pubs les jours suivants. Si les choses s'arrêtent là, on ne peut pas franchement considérer qu'il y a intrusion.

Mais les choses peuvent ne pas s'arrêter là. Car les données que nous venons de décrire, vos achats et vos consultations chez Darty, il se peut que la machine qui va les gérer ne soit plus à Darty, ou que Darty en sous traite le travail à une entreprise - que j'appelle

Alpha - qui gère également le commerce en ligne de Digitick (places de spectacle, de musée, etc.) et d'autres encore. A ce moment-là, Alpha dispose d'informations nouvelles, dans différents domaines, et qu'il peut croiser. Mais là encore, ce n'est pas vraiment l'individu qui va à la fois prendre son ticket d'entrée au musée et qui achète chez Darty qui l'intéresse.

La machine va plutôt être programmée pour savoir si, statistiquement, les gens qui vont au Musée achètent telle gamme de produits bon marché ou s'ils sont plutôt prêts à se payer de meilleurs produits. On peut aussi chercher si ces personnes seraient enclines à acheter des produits high-tech, nouveaux, si l'âge entre en ligne de compte, etc. La machine va être programmée pour trouver des corrélations, du genre : les jeunes qui vont au musée se paient des produits nouveaux et de qualité ; les vieux qui vont au musée, au contraire, restent dans le traditionnel. Grâce à quoi, lorsqu'une personne ira sur le site de Digitick, la machine lui enverra immédiatement le type de publicités que l'on pense lui correspondre.

Encore une fois, et jusqu'ici tout au moins, ce n'est jamais vous en tant qu'individu qui intéresse la machine qui récupère vos données. Vous n'êtes qu'un élément d'une grande statistique. Tout le monde de la publicité en ligne, sur internet, fonctionne ainsi. Cela représente un chiffre d'affaires de près de 5 milliards d'euros pour 2018. Nos données intéressent le secteur, et elles ont donc de la valeur. Lorsque nous recevons des publicités par la poste, si ce n'est pas nous qui avons donné notre adresse à l'expéditeur, c'est qu'elle lui a été vendue, ou revendue. Et c'est la même chose avec nos adresses mail. Nos adresses, nos données sont donc des marchandises, qui vont et viennent.

L'avocat François Sureau, grand défenseur des libertés démocratiques dans le système bourgeois, l'a bien compris. Il le dit dans son opuscule "Sans la liberté" : *"Ces grandes sociétés n'obéissent pas aux mêmes déterminants que les États classiques. Leur but n'est pas, a priori, de contrôler les comportements (...) mais seulement de les connaître pour vendre davantage et enrichir leurs actionnaires. Nul ne peut cependant exclure que ceux qui les dirigent se découvrent animés d'une soif de puissance de nature politique, ne forment de vastes desseins, n'utilisent à cette fin les données dont ils disposent. Mais quand bien même ce ne serait pas le cas, c'est, du point de vue du citoyen, la collusion entre ces nouveaux empires et les acteurs étatiques traditionnels qui est à redouter (...)"*.

Effectivement, si Darty, ou Digitick ne nous fliquent pas vraiment, individuellement, quand ils nous envoient des pubs, cela ne veut pas dire que ce n'est pas possible. Puisque tout s'achète et tout se vend, il est possible, si on en a les moyens, d'acheter suffisamment de données pour opérer si ce n'est un flicage, du moins une opération d'influence politique de masse. C'est ce qu'a fait Cambridge Analytica.

En 2014 et 2015, un professeur de psychologie de l'université de Cambridge, Aleksander Kogan, lance un quiz qu'il appelle "thisisyourdigitallife" (ceci est votre vie numérique), qu'il présente comme une étude scientifique. Le quiz est téléchargé par 277 000 utilisateurs de Facebook, qu'on paye 4 dollars pour répondre aux questions du quiz. Mais les données, et aussi celles de tous leurs amis sur Facebook, sont récupérées, sans bien sûr que les gens en soient informés. Cela va concerner 87 millions de personnes (dont 210 000 Français). Grâce à quoi, donc, la société Cambridge Analytica aurait tenté d'influencer le vote des électeurs, en 2016, pour favoriser l'élection de Trump.

Une enquête est en cours. Mais on voit d'ores et déjà qu'une société qui parvient à stocker autant d'informations peut se prévaloir d'être utile à un candidat, que ce candidat aura

donc intérêt à faire appel à ses services, donc à la financer, c'est-à-dire à lui donner encore plus de moyens pour qu'elle aspire de nouvelles données. Tout cela n'est donc finalement qu'une question d'argent, ce qui nous indique que nous sommes toujours au cœur d'une société bien capitaliste.

Des données qui peuvent être capturées, puis analysées, il y en a de plus en plus dans notre vie quotidienne, d'autant que tout est organisé pour nous inciter, voire nous obliger, à passer par internet. Une image postée sur un réseau social, un rythme cardiaque mesuré avec une montre connectée, le nombre de pas que compte le podomètre sur votre smartphone, votre positionnement, votre emplacement, les lieux que vous demandez à votre moteur de recherche de trouver, un travail partagé avec des collègues sur le net, des milliards de mots, d'images, d'horaires, de localisations, etc. s'amassent chaque seconde sur les ordinateurs.

Pour l'informaticien Antoine Petit, *“le vrai pouvoir est de rassembler énormément de données, ce qu'ont bien compris les grandes plates-formes numériques, qu'elles soient américaines ou asiatiques”*. Il suffit de regarder les places en Bourse des plus grandes sociétés actuelles : six des dix premières possèdent des collectes massives de données : les américaines Amazon, Apple, Alphabet-Google, Facebook, et les chinoises Alibaba (plus grande plate-forme de commercer en ligne au monde) et Tencent (services en ligne : jeux, réseaux sociaux, messagerie, publicité, commerce). Et si aux États-Unis, on a les GAFAs (Google, Apple, Facebook, Amazon), en Chine, nous avons les BATX (Baidu, Alibaba, Tencent et Xiaomi).

Pour David Chavalarias, directeur de recherche au Centre d'analyse et de mathématiques sociales, *“ces types d'acteurs collectent désormais des interactions sociales avec une granularité extrêmement fine sur des millions d'utilisateurs. Ils peuvent tout à fait connaître chacun de leurs utilisateurs dans le détail et cerner des dynamiques sociales globales. Cette connaissance est monnayée au sein de leurs régies publicitaires”*.

Et là, c'est clair, nous ne sommes plus dans la simple publicité qui nous choque. Là, dans ces grands groupes, on peut effectivement cliquer l'individu que nous sommes. On peut choisir une personne et reconstituer si ce n'est toute sa vie, du moins une grande partie de son activité.

Une fois ces montagnes de données collectées, il n'y a plus qu'à concevoir des programmes, des algorithmes, pour fouiller dedans, chercher ce qu'il est possible d'en conclure. Il s'agit donc, à ce stade, de faire parler le stock d'informations qu'on a accumulées. Et on assiste actuellement à une course mondiale pour recruter, au prix fort, les meilleurs scientifiques, les meilleurs informaticiens, les meilleurs mathématiciens. Des salaires cinq ou dix fois plus hauts que ceux du public leurs sont offerts. Et c'est une forme de fuite des cerveaux qui a lieu, évidemment, toujours vers les mêmes plus grosses entreprises, les GAFAs (Google, Apple, Facebook, Amazon) pour ce qui est des USA.

Les laboratoires publics reçoivent de plus en plus de financement des GAFAs. Et celles-ci leur proposent un accès à telle ou telle partie de leurs bases de données. Cela aide, certes, la recherche publique. Mais dans ces conditions, il y a peu de chance que des travaux de réflexion critique y soient menés. Ces liens ne peuvent qu'inciter les laboratoires publics à orienter leurs recherches dans un sens particulier qui intéresse les GAFAs.

La société privée américaine 23andMe (filiale de Google) vend des kits qui vous permettent d'analyser votre ADN. Avec votre accord, et il est le plus souvent donné, elle garde ensuite une copie de cette analyse. Résultat, c'est elle aujourd'hui qui possède la plus grande collection au monde de génomes humains.

L'IA et la médecine

Les exemples de réussite de l'IA sont déjà nombreux dans le domaine médical. Le 11 avril 2018, la FDA, l'Agence américaine des médicaments, a autorisé l'usage d'un logiciel IDx-DR, qui est le premier système de diagnostic autonome basé sur l'intelligence artificielle à être autorisé sur le marché. Ce système est destiné à détecter des anomalies sur des photographies de la rétine. Il faut savoir que les diabétiques sont sujets à une maladie de la rétine, la rétinopathie, et que celle-ci ne pouvait, jusqu'ici, se déceler que par un examen minutieux des images ; il faut déceler des taches hémorragiques, une dilatation ou une croissance anormale des vaisseaux sanguins, et cela prend, au total, des milliers d'heures d'examen aux ophtalmologues. La machine, elle, entraînée - vous l'avez deviné - par l'ingurgitation de milliers et de milliers d'images, va réussir ce travail : elle identifie la maladie dans 87% des cas où il fallait le faire, et dit qu'il n'y a pas de maladie dans 89,5 % des cas où effectivement il n'y en avait pas, ce qui est aussi bien ou mieux qu'un ophtalmologue.

Libération hors-série sur l'IA qui donne cette information s'extasie : *“Le potentiel est déjà vertigineux : rétinographies, mais aussi échographies, IRM, scanners ou simples radios des os... Un algorithme nommé OsteoDetect est ainsi spécialisé dans les poignets cassés : il marque d'une flèche l'emplacement d'une fracture s'il en détecte une sur la radio. Le logiciel VizLVO, lui, récupère les occlusions d'artères dans les scanners cérébraux : dans un hôpital où l'équipe de nuit est réduite, il accélère la prise en charge des AVC en convoquant directement par texto le neurologue interventionnel d'astreinte, qui visualise le scan sur son smartphone et déclenche l'opération. Tous deux ont également reçu un feu vert de la FDA cette année (on est fin 2018). Et derrière ces précurseurs, des dizaines de start-up font la course pour être les prochaines sur la liste.”*

Pour mettre au point tous ces logiciels, on ne le répétera jamais assez, il faut donner à avaler à la machine des tonnes d'images médicales. Et c'est donc la course pour les trouver, partout sur la planète. En France, la chose est un peu plus facile. Du fait d'une tradition de centralisation, il suffit de s'adresser à l'AP-HP, qui a créé pour cela un Entrepôt de données de santé (EDS, autorisé par la CNIL en janvier 2017), pour favoriser leur utilisation à grande envergure dans les projets de recherche et de mise au point des algorithmes. L'EDS rassemble 10 millions de patients. Seuls quelques centaines s'y sont opposées, mais il est probable que la plupart de ceux qui ont accepté n'en ont guère eu conscience, tant la formalité est intégrée dans les divers formulaires administratifs.

Tout est stocké dans l'EDS : *“les données démographiques, les archives des prises en charge hospitalières, les diagnostics, les actes médicaux, les résultats d'analyses biologiques, les examens d'imagerie avec leur compte-rendu”*. On peut donc aller chercher, pour étude, ce que l'on veut. Par exemple, les femmes qui font une fausse couche récidivante entre 35 et 45 ans, ou les patients qui ont dû être hospitalisés pour une grippe entre 2011 et 2013, ou encore 300 000 images de mammographie.

Depuis une loi de janvier 2016, la base de données de l'Assurance maladie, déjà largement ouverte aux chercheurs du public, l'est aussi, sous certaines conditions, aux chercheurs du privé.

Dans la course au ramassage des données, on trouve en excellente position... La Poste. Selon Libération (déjà cité) elle dispose d'une filiale Docapost. Après avoir signé un accord avec l'ordre des pharmaciens, elle aspire tous les dossiers que vous voyez remplir chez votre pharmacien, et elle héberge ainsi les données de santé de 40 millions de Français. On sait que la Poste propose déjà un service de visite à domicile en milieu rural, pour les personnes âgées. On peut penser qu'elle a d'autres idées en tête encore.

L'autre grande idée après les algorithmes, concernant la santé, c'est la connexion. Il y a actuellement 6,1 millions de personnes âgées de plus de 75 ans en France. Sur ce chiffre, il y en a 720 000 dans des établissements spécialisés. Or l'on sait que les problèmes de santé augmentent considérablement avec l'âge. On envisage donc des solutions de maintien à domicile, avec par exemple des consultations médicales virtuelles. Un infirmier qui rend visite à la personne peut être équipé d'un stéthoscope connecté et d'une tablette. Il peut ainsi faire écouter la respiration ou le rythme cardiaque par un médecin à distance. Depuis le 15 septembre 2018, les téléconsultations sont ainsi autorisées en France.

On peut aussi équiper les personnes isolées d'un bracelet connecté, qui peut enregistrer la température, le rythme cardiaque, le nombre de pas, voire une chute, avec un bouton d'alarme. Une filiale du Crédit Mutuel-Arkès propose une montre avec alerte vidéo et géolocalisation. La société spécialisée dans les sols Gelflor propose, de son côté, un système de capteurs à installer sous le parquet ou le lino, dans les maisons médicalisées : le champ électrique sera impacté en cas de chute, et enverra automatiquement un sms ou un mail d'alerte ; coût 3500 euros le mètre carré. On le voit, les idées ne manquent pas, et le marché est avide de trouver de nouveaux débouchés.

Nous sommes dans un système capitaliste. Ce qui veut dire que toute nouveauté qui semble efficace, prometteuse d'avenir, va immédiatement susciter une course pour en profiter et devenir un des gagnants. Et dans cette course, beaucoup de coups sont possibles. Certains peuvent bluffer, tenir des discours enflammés, annonçant des prodiges, en vue d'attirer les subventions de l'Etat, les prêts des banques, les capitaux des investisseurs. D'autres s'enflamment parce qu'ils y croient vraiment. Il est difficile de faire le tri.

Sur l'intelligence forte

Il y a une idée, une question, qu'il est légitime de se poser, à propos de l'intelligence artificielle. Nous avons dit et souligné qu'actuellement, chaque IA n'est utilisable que dans son tout petit domaine. Et que de ce fait, on appelle cette forme actuelle de l'IA une intelligence faible. Certains l'opposent à une éventuelle, ou prochaine, intelligence forte.

Une intelligence forte serait une intelligence capable de cerner des domaines différents, une intelligence capable de tenir compte de tout l'environnement, au lieu de se cantonner à une tâche unique, sans avoir aucune idée du contexte, de ce qui peut se passer autour. On ne peut pas exclure que, dans une décennie ou deux, ou peu importe quand, la recherche et les progrès parviennent à avancer dans cette voie.

Selon Jean-Gabriel Ganascia, professeur d'informatique à Sorbonne Université, un certain nombre de chercheurs engagés dans l'intelligence artificielle affirment que l'on peut parvenir à une IA forte (Strong artificial intelligence) *“qui reproduirait un esprit, voire une conscience, sur une machine. On assure même que cette dernière surviendra très bientôt, que c'est inéluctable, que ce sera irréversible et que cela aura des répercussions majeures, tout à la fois positives et négatives, sur le devenir de l'espèce humaine”*. (La Vie). Le moment fatal où l'on basculerait dans cette nouvelle situation, on lui a donné un nom : *“la singularité technologique”*.

La "singularité", c'est donc le *“moment où la machine, capable de s'autoperfectionner jusqu'à créer une superintelligence, deviendrait plus puissante que notre cerveau. Elle détroterait alors l'homme et signifierait, selon l'écrivain américain Vernor Vinge, la fin de l'ère humaine”* (La Vie).

Depuis les années 2000, il existe des équipes de chercheurs qui souhaitent arriver à ce stade, et qui tentent d'y travailler d'arrache-pied. On appelle ce courant l'intelligence artificielle générale. Ses partisans sont tout à fait conscients des limites de l'IA faible. Selon Jean-Gabriel Ganascia, *“ses promoteurs désirent refonder l'IA sur des bases mathématiques solides, équivalentes en certitudes à celles sur lesquelles s'appuie la physique. Dans ce but, ils recherchent une formule clé qui résumerait le principe général sur lequel reposeraient toutes les formes d'intelligence. À titre d'illustration, certains s'appuient sur un principe de minimisation de la quantité d'informations, autrement dit d'une généralisation ultime des données d'expérience qui retiendrait l'essentiel. Pour cela, ils ont recours à des théories mathématiques d'optimisation (...)”*

Je vous épargne la suite. Quoi qu'il en soit, Jean-Gabriel Ganascia conclut : *“le danger ou l'espoir que représentent l'IA forte apparaissent plus imaginaires que réels. Et il en va de même du risque de voir advenir la singularité technologique”*.

Voici également la réponse, qui nous semble matérialiste, que donne à cette interrogation Serge Tisseron, psychiatre, directeur de recherche à l'université Paris-Ouest-Nanterre : *“En supposant qu'un jour, la machine ait une conscience, nous dit-il, elle la fabriquera nécessairement à partir des capteurs qui l'informeront sur le monde. Or ces capteurs seront différents de nos yeux, de nos oreilles, de notre peau. Par exemple, ses "yeux" seront sensibles à toutes les longueurs d'ondes, ses "oreilles" capteront une large gamme de fréquences... la machine aura alors sur le monde des informations différentes de celles dont nous disposons. Avec une telle expérience du monde, elle aura une conscience différente de la nôtre. Sa conscience ne sera donc en aucun cas une superconscience humaine ! Ce serait une erreur grave de penser qu'elle puisse nous éclairer sur notre propre expérience du monde. Notre conscience est fondée sur nos corps de chair et de sang façonnés par des millions d'années d'évolution, celle de la machine résultera d'un assemblage de métaux et de plastique... Une telle intelligence artificielle devrait donc faire de grands efforts pour comprendre notre perception du monde, et il n'est pas certain qu'elle en ait envie”*. (Pour la Science hors-série n°87 avril 2015)

Les masques du transhumanisme

En 2015, un appel signé par 1000 personnes a été lancé pour dénoncer les dangers de l'intelligence artificielle. Il dénonçait les risques dus à un développement des machines et de

leurs capacités. Or, on trouve parmi les signataires de nombreuses personnalités importantes du monde de l'IA. Il y a des chercheurs en robotique, il y a un Elon Musk (Tesla, Space X et les puces du cerveau), un Steve Wozniak, cofondateur de Apple, il y a un Dennis Hassabis, fondateur de DeepMind, l'entreprise d'intelligence artificielle appartenant à Google. Il y a aussi des gens comme Stephen Hawking (les trous noirs), Noam Chomsky (linguiste de l'extrême gauche américaine).

Chomsky ou Hawking ne connaissent rien à l'intelligence artificielle. Mais les trois premiers cités ici sont au cœur même du monde actuel de l'IA. Leur message, c'est *“On a peur des machines qui auront des compétences statistiques et combinatoires supérieures aux nôtres, ainsi qu'un accès à des bases de données gigantesques que nous, humains, ne pouvons pas manipuler”*.

Pourquoi signent-ils, ou peut-être même ont-ils initié, cet appel ? On a une réponse à cette question dans Pour la Science (hors-série n°87 avril 2015). Serge Tisseron et Frédéric Alexandre sont, le premier psychiatre, directeur de recherche à Paris-Ouest-Nanterre, le second informaticien à l'Inria Bordeaux. Pour eux, cet appel est une manière de noyer le poisson, de détourner l'attention d'une manière ou d'une autre sur les changements que l'IA actuelle est déjà en train de procéder.

Les gens qui ont lancé cet appel crient bien fort sur le risque d'avoir affaire à une intelligence forte, alors qu'elle n'existe pas, qu'elle n'est pas près d'exister. Frédéric Alexandre explique : *“on pourrait parler d'enfumage ! (...) Donnons un exemple. Certains logiciels utilisés dans la finance, notamment celle à haute fréquence, procèdent à des mouvements de capitaux, spéculent, prennent des positions en une fraction de seconde. Ces programmes sont beaucoup plus rapides que l'intelligence humaine et lui sont donc supérieurs en termes de puissance combinatoire et mathématique. La vraie question à se poser n'est pas tant sur leurs performances que sur les valeurs qu'on leur a inculquées. Leur but est-il de faire le plus d'argent possible à court terme ? Ou bien prennent-ils en compte d'autres critères, par exemple le respect de l'économie réelle, le développement agricole des pays pauvres ? Le vrai problème est là. Et, pour le moment, ce sont bien des êtres humains qui définissent les objectifs de ces logiciels.”*

Autrement dit, la capacité des machines à aller très vite, ou à utiliser des montagnes de données, ce n'est pas cela qui est dangereux. Ce qui peut être dangereux c'est ce qu'on a dit à la machine de faire, et qu'on a traduit en algorithmes. Et cela, ce sont des humains qui le font.

“Quand on regarde avec attention les auteurs de la lettre ouverte, explique Frédéric Alexandre, on observe que la plupart sont impliqués dans les grandes sociétés qui, justement, commercialisent des produits de ce type. Au lieu de pointer les risques de la combinatoire, ils feraient mieux de rendre publiques les valeurs qu'ils intègrent dans leurs logiciels. Quels est la place de l'humain ? La place de la finance ? La place de la spéculation à court terme ? L'enjeu est là.”

Des enfumeurs, ou des gens qui se font avoir par cet enfumage, on en trouve beaucoup à la lecture de tout ce qui s'écrit actuellement sur l'IA. En voici un autre, interviewé par Finkielkraut sur France Culture, Laurent Alexandre (à ne pas confondre avec Frédéric Alexandre, que l'on vient de citer). Chirurgien urologue, entrepreneur, Laurent Alexandre est présenté dans Libération comme *“l'ambassadeur des nouvelles technologies”*. Très fort

partisan du développement de l'IA, il tient un langage qui semble celui de la prudence, mais nous allons voir qu'il s'agit d'une stratégie. *"Avoir une intelligence artificielle qui raisonne, nous dit-il, pose le problème de son hostilité éventuelle. Nous ne savons pas encore si une intelligence consciente d'elle-même cherche forcément à exterminer les autres - comme nous avons tué l'homme de Néandertal dans notre évolution il y a trente mille ans - ou s'il peut y avoir des intelligences qui ne sont pas hostiles, qui cohabitent avec les autres sans chercher à prendre le leadership. Nous ne le savons pas bien. Dans le doute il faut faire attention"*.

Et il en arrive à cette conclusion : *"S'il devait arriver un jour qu'on ait des intelligences fortes, il faudrait sans doute une police de l'intelligence artificielle pour éviter que la science-fiction ne devienne réalité. Mais on en est loin aujourd'hui."*

En clair, et en décodant : *"l'intelligence forte, ce sera dangereux, mais ce n'est pas pour demain."* Et on en conclut implicitement que l'intelligence actuelle, faible, ça n'est pas dangereux, et il n'y a pas à s'en inquiéter, pas besoin, comme il dit, de "police".

Mais si on fouille un peu dans la manière de voir de ce monsieur Laurent Alexandre, on trouve vite de quoi nous inquiéter. Il a écrit un livre intitulé *La Guerre des intelligences* où on lit ce genre de choses : *"L'intelligence n'est qu'une question de quantité d'opérations traitées, autrement dit de puissance de calcul. (...) Le cerveau est un ordinateur fait de viande, extraordinairement sophistiqué."* Il propose d'abandonner le QI (quotient intellectuel) pour passer au QCIA, un *"quotient de complémentarité à l'intelligence artificielle"*, qui semble bien vouloir faire du cerveau humain un annexe de la machine.

Quant à l'avenir, voilà comment il le voit : *"À partir de 2030 environ, l'éducation va sortir de l'âge du bricolage pour devenir une science exacte. Demain, l'apprentissage sera une technologie (...). L'école deviendra transhumaniste et trouvera normal de modifier le cerveau des élèves en utilisant toute la panoplie des technologies NBIC"* (cet acronyme signifie : nanotechnologies, biotechnologies, informatique et sciences cognitives). Et il continue : *"Le gros du travail aura lieu avant même l'entrée à l'école proprement dite, la sélection embryonnaire sera réalisée lors de la conception et la neuro-augmentation électronique, sur les très jeunes enfants. Le travail de l'équipe éducative débutera avant la naissance puisqu'elle aidera les parents à paramétrer la sélection embryonnaire."*

Dans le numéro 170 de fin 2019 de la revue Pouvoirs consacré à l'Intelligence artificielle, où s'expriment aussi Cedric Villani et Luc Ferry, le même Laurent Alexandre prétend résoudre les inégalités sociales grâce à l'intelligence artificielle et au transhumanisme. Il explique que les inégalités, l'école ne parvient pas à les réduire, ce qui est vrai. Mais il continue en affirmant que l'erreur est que les éducateurs, les enseignants, font comme si tout le monde avait le même cerveau à la naissance. Selon lui, et d'autres sur lesquels il s'appuie, c'est faux. *"L'ADN détermine plus de 50% de notre intelligence"* prétend-il. Et *"ce n'est pas parce qu'il y a des livres dans les bibliothèques des bourgeois que leurs enfants sont de bons lecteurs : c'est parce qu'ils ont reçu un bon patrimoine génétique. 64% de nos différences en matière de capacité de lecture sont d'origine génétique : la famille, l'école, nos efforts individuels n'en expliquent qu'un tiers. (...) Comme l'explique Franck Ramus, spécialiste de la cognition à l'École normale supérieure : "En moyenne, les personnes les plus défavorisées socialement sont aussi les plus désavantagées génétiquement."* - Vous avez bien entendu, il semble qu'on puisse trouver ce genre de discours à Normale sup ! -.

Mais où Laurent Alexandre veut-il en venir ? Eh bien, il conclut de tout ceci qu'il n'y a qu'à changer les cerveaux de ceux à qui la nature et l'ADN n'en ont pas donné un assez bon. Et il propose pour cela deux méthodes : la sélection, en analysant l'ADN avant la naissance, et, on le suppose, en manipulant l'ADN de ceux dont on pense qu'il ne développerait pas un assez bon cerveau. Ou par une greffe de puces sur le cerveau, pardon par l'"augmentation cérébrale".

Et d'applaudir Elon Musk qui a annoncé, le 28 mars 2017, *“le lancement de Neuralink, une société destinée à augmenter nos capacités cérébrales grâce à de minuscules composants électroniques entrelacés à nos quatre-vingt-six milliards de neurones, ce qui nous transformerait en cyborgs. Elon Musk se donne cinq ans pour sortir les premiers prototypes.”* Et d'ajouter, pour rendre aussi crédible son envie de manipuler l'ADN, *“Par ailleurs, les premiers bébés génétiquement modifiés ont été fabriqués en Chine par He Jiankui, un chercheur à l'université de Shenzhen. (...) Il ne fait pas de doute que l'on va pouvoir, à terme, augmenter le QI des bébés par de telles manipulations”*.

On peut hausser les épaules face à ces discours. Sauf que monsieur Elon Musk, milliardaire fondateur de Space X (les fusées qui se reposent sur Terre, de Tesla (automobile électrique), propriétaire un moment de Paypal (moyen de paiement sur internet) a réellement investi son argent sur ces recherches visant à "augmenter" - c'est son terme - l'humain et son cerveau.

À un interlocuteur qui dit son horreur à envisager un tel monde, Laurent Alexandre répond avec cet argument bien écœurant : *“Il sera bien difficile (quand ces microprocesseurs existeront) d'expliquer au paysan tanzanien pauvre que c'est très mal de mettre les microprocesseurs d'Elon Musk dans le cerveau de ses enfants pour qu'ils fassent Harvard et qu'il vaut mieux qu'ils restent des paysans miséreux”*.

Pour conclure sur la stratégie des GAFAs, il est important de comprendre leur argumentation. Ils commencent à nous dire : *Regardez ce qu'on fait avec l'IA, c'est nouveau, c'est formidable. Faites-nous confiance.* Mais dès que quelqu'un réfléchit, se pose des questions, et signale qu'il ne faut pas aller n'importe où, la réponse, c'est : *Mais ça existe déjà, c'est déjà voulu par les gens. C'est dans la suite logique de ce qui s'est toujours fait. C'est le progrès. Ça va se passer de toute manière. Pas la peine de discuter !*

Eh bien, si, il faut discuter, du monde, de son avenir, et de la société. En attendant, les principaux dirigeants des GAFAs, eux, ont fait le choix de mettre leurs enfants non pas dans des écoles qui investissent à tour de bras dans l'intelligence artificielle, mais dans une école... déconnectée. Cela nous rappelle la fameuse réponse du PDG de Marlboro, qui expliquait qu'il gardait sa "merde" pour les immigrés, les Noirs ou les pauvres.

Dans un numéro du Monde, Laurent Alexandre, toujours lui, applaudissait déjà à cette nouvelle le 18 avril 2013 : *“Ray Kurzweil, le "pape" du transhumanisme, a été embauché par Google comme ingénieur en chef pour faire du moteur de recherche la première intelligence artificielle de l'histoire”*. Il écrivait avec enthousiasme : *“Aujourd'hui, Google est devenu l'un des principaux architectes de la révolution NBIC et soutient activement le transhumanisme, notamment en parrainant la Singularity University qui forme les spécialistes des NBIC. Le terme Singularity désigne le moment où l'esprit humain sera dépassé par l'intelligence artificielle, censée croître exponentiellement dès les années 2045. Ray Kurzweil, le "pape" du transhumanisme, dirige en personne cette université. Ce*

spécialiste de l'intelligence artificielle est convaincu que les NBIC vont permettre de faire reculer la mort de façon spectaculaire dès le XXI^e siècle. Il a été embauché par Google comme ingénieur en chef pour faire du moteur de recherche la première intelligence artificielle de l'histoire.

“Google s'intéresse également au séquençage ADN au travers de sa filiale 23andMe, dirigée par la femme de Sergei Brin, le cofondateur de Google. Sergei Brin a appris qu'il avait de très fortes chances de développer la maladie de Parkinson - car il est porteur de la version mutée du gène LRRK2 - en faisant analyser son ADN par sa filiale. De quoi accentuer son intérêt pour les NBIC ! On aura compris que Google nous emmène vers une civilisation transhumaniste.”

À notre connaissance, il n'y a encore nulle part au monde de cerveau humain "transhumanisé". Mais ce courant pousse ses pions partout où il le peut. En Suède, plusieurs milliers de personnes ont accepté de se faire greffer dans la main, entre le pouce et l'index, un implant, avec toujours cette idée de faire fusionner l'homme et la machine. L'implant ayant enregistré leur identité, ils se vantent de ne plus avoir besoin de taper leur code personnel dans les bureaux où ils travaillent pour avoir accès au photocopieur, ou à payer chaque consommation à la cafétéria. *“On entre dans une nouvelle ère, se vante Hannes Sjöblad, le promoteur de ces implants NFC, dans laquelle la technologie migre à l'intérieur du corps. On finira par créer des implants actifs, qui seront capables de déchiffrer ce qui se produit dans nos corps, et de le transmettre”.*

Ces robots qui fascinent les humains

À l'heure actuelle, il existe un robot, appelé Paro, qui a l'apparence d'un phoque en peluche, et qui est utilisé dans des maisons médicalisées d'une trentaine de pays. Il y en a 40 en France, 3000 dans le monde (chiffres donnés par Laurence Devillers en 2017). Le robot Paro a d'abord été commercialisé au Japon en 2005, puis aux États-Unis en 2009, où il a été autorisé par la FDA.

Paro aide à la fois les patients, ceux atteints de maladies du type Alzheimer, et les soignants. Il comporte une douzaine de capteurs, qui enregistrent le toucher, la lumière, le positionnement, et trois microphones. Il répond au toucher par de petits cris. Selon les informations qu'il reçoit ainsi, son algorithme adapte ses mouvements en conséquence, son intonation. Il peut communiquer au patient ce que l'humain considère comme une émotion, telle que la joie ou la colère. Il crée une atmosphère de compagnie, apaisante, reposante.

Dans la pratique, sa présence incite les patients à un transfert de sentiments, dans certains cas à une réminiscence de souvenirs, et provoque surtout une baisse de l'anxiété, ce qui permet une diminution de l'utilisation des médicaments. Il est vendu autour de 5 000 euros.

Pour Laurence Devillers, professeure en informatique appliquée à Paris-Sorbonne, qui décrit Paro dans son livre *Des robots et des hommes*, ce type de robot est tout à fait utile. *“Des robots plus évolués comme les Marvin permettront de maintenir les personnes âgées ou dépendantes plus longtemps chez elles. Les robots seront là sept jours sur sept, jamais fatigués, pour surveiller et assister les patients à domicile, en respectant la dignité des humains, sans jugement et en pouvant répéter les consignes et gestes sans jamais montrer*

d'impatience. Ils seront utiles : pour éviter la solitude, l'anxiété, les détresses, voire au pire des envies suicidaires ; pour aider les personnes dans leur intimité et préserver leur dignité : dans la toilette et l'habillement ; pour éviter les chutes, pour aider dans les gestes du quotidien ; pour le ménage et le soin d'animaux de compagnie ; et pour prévenir le médecin et les proches en cas de problèmes”.

Les robots existent déjà. Et comme pour ce que nous avons vu de manière générale avec l'intelligence artificielle, certains ont voulu utiliser cette existence pour agiter, ou promouvoir, un avenir qui fasse basculer l'homme par la machine. Ainsi, des voix ont proposé qu'on réfléchisse au statut des robots. Comme on a reconnu en France à l'animal le statut d'être sensible, avec des sanctions à la clé, il a été proposé que, d'ores et déjà, on reconnaisse aux robots actuels le statut d'être sensible. Puisqu'on arrivera un jour à la fameuse "singularité", et que les robots seront alors supérieurs à nous, on devrait considérer ceux actuels comme des petits êtres en devenir qui méritent le même statut juridique.

Cette argumentation a, un moment quand même, touché jusqu'aux autorités européennes en matière de rédaction du droit et, dans un premier projet, on a bel et bien écrit qu'une conscience de robot pourrait apparaître. Finalement, cette idée a été effacée de la résolution du 12 février 2019 sur la politique industrielle européenne globale sur l'intelligence artificielle.

Un écrivain américain de science-fiction, Isaac Asimov, s'était très sérieusement posé des questions concernant les robots. Il en avait conclu qu'il serait nécessaire, dans un monde où les robots seraient présents, de décider de trois lois, à traduire dans la manière de les fabriquer et de les programmer. Eh bien, le projet de rapport de 2016 de la Commission européenne sur les règles de droit civil en robotique, a décidé qu'*“il y a lieu de considérer les lois d'Asimov comme s'appliquant aux concepteurs, aux fabricants et aux utilisateurs de robots”*.

Ces lois, rédigées en 1942, on les trouve dans la nouvelle *Runaround* ("cercle vicieux" ou "cycle fermé", selon la traduction d'Asimov) :

- 1- Un robot ne peut attenter à la sécurité d'un être humain, ni, par inaction, permettre qu'un être humain soit mis en danger.
- 2- Un robot doit obéir aux ordres d'un être humain, sauf si ces ordres entrent en conflit avec la première loi.
- 3- Un robot doit protéger sa propre existence tant que cela n'entre pas en conflit avec la première ou la deuxième loi.

Un peu plus tard, Asimov jugera nécessaire d'ajouter une quatrième loi, qu'il a appelé loi 0 :

0- un robot ne peut ni porter atteinte à l'humanité, ni permettre que l'humanité soit exposée à un danger.

Ce qui est ici à retenir, c'est que pour la loi européenne, c'est bien l'humain qui contrôle la machine, c'est donc lui qui est responsable.

L'Arabie saoudite veut construire un nouveau type de ville, qui serait pensé pour être peuplée par plus de robots que d'humains. Les Émirats arabes unis, de leur côté, envisagent de se doter d'une police formée à 25% de robots, d'ici 2030. Les robots vont-ils nous envahir ? Là encore, on peut se perdre en spéculations à perte de vue.

Nous nous intéresserons plutôt à un aspect de la relation homme-robot qu'étudie Laurence Devillers. Elle insiste sur cette fascination que ressentent les humains face à un robot qui leur ressemble plus ou moins. Effectivement, si votre robot aspirateur n'est qu'une machine aplatie qui circule seule, enregistre le plan de ce qui se trouve sur le sol de l'appartement, et si besoin l'actualise, vous trouvez cela plutôt ingénieux et pratique. C'est une machine. Mais la même machine, si on lui donne un aspect vaguement humain, produirait sur nous un tout autre effet.

“Le robot en Occident, explique Laurence Devillers, nous a longtemps effrayé car il était supposé malintentionné. Au Japon, on considère que les objets, le vent, les arbres, les pierres, etc. sont animés de vie et protecteurs.” Peut-être aussi la tradition judéo-chrétienne tend-elle à condamner la fabrication de machines pouvant ressembler à l'homme, la vie étant du domaine réservé de Dieu. En tout cas, pour les humains, il peut y avoir un attrait pour les robots, ou une répulsion, rarement une indifférence. Mais du coup, constate Devillers, on peut assez vite en arriver à ce que se crée une empathie entre un être humain et son robot, dès lors que celui-ci est régulièrement à son service.

“La machine n'a pas d'émotions, ni de sentiments, ni de conscience en l'état actuel de nos processeurs”, insiste-t-elle. N'empêche, elle est programmée pour détecter les émotions chez l'humain, comme dans l'exemple de Paro. Et cela suffit pour que l'humain, presque instinctivement, commence à voir son robot comme quelqu'un qui le comprend. Il peut aussi rapidement se mettre à aimer sa machine. *“Certains donnent des noms à leurs robots aspirateurs, nous dit-elle, et si le robot tombe en panne veulent qu'il soit réparé et non échangé contre un neuf car il connaît leur maison, preuve que l'humain projette une identité sur le robot”.*

Devillers travaille donc à essayer de prévenir les risques liés à l'utilisation des robots, risques qui proviennent de la manière dont les humains les voient : *“L'addiction et l'isolement, ainsi que le report d'autonomie sur la machine, la confusion entre la machine et l'humain sont des déviations dont il faut s'occuper. (...) Un autre risque est d'oublier qu'un robot ne ressent rien, n'a pas d'émotion, n'a pas de conscience et n'est pas vivant. Il est possible de ressentir de l'empathie pour un robot (...). Il est important que les personnes âgées qui peuvent mettre leur vie en danger pour venir en aide à leur robot se rendent compte qu'un robot ne souffre pas même s'il tombe, il faut qu'elles soient conscientes que ce n'est qu'un objet programmé...”* Dans le domaine militaire, on se soucie aussi que le soldat ne s'attachera pas trop à un robot qui le soutient, au point de risquer peut-être sa vie pour la machine...

Quel avenir pour le monde du travail ?

L'utilisation de plus en plus développée de l'IA va-t-elle supprimer de l'emploi de façon importante, va-t-elle en créer de nouveaux ? Ou peut-être avec un peu de chance les choses seront-elles finalement équilibrées entre les destructions d'anciens emplois et la création de nouveaux ? On trouve autant de réponses que l'on veut à ces questions.

Dans le petit opuscule intitulé - de manière un peu alléchante - *L'intelligence artificielle, de la révolution technologique à la révolution sociale*, Nicolas Spatola, qui travaille sur les interactions homme-machine à l'université de Clermont, pose la question angoissante : *“Les intelligences artificielles sont-elles amenées à remplacer les emplois ?”* Il

note que ce ne serait pas la première fois : *“Dans l'agriculture, on observe par exemple un déclin de 95% des emplois entre 1871 et aujourd'hui”*. Mais aujourd'hui, pense-t-il *“en raison des limites de dextérité et de capacité cognitive des robots, les métiers les plus susceptibles d'être remplacés sont ceux représentant une routine ou ne requérant pas une grande dextérité manuelle, à l'instar de la comptabilité ou du support client”* Il y aurait aussi ces domaines qui nécessitent une gestion de données à grande échelle : la partie diagnostic du métier de médecin, la jurisprudence pour les avocats. Quant aux fameux camions ou automobiles autonomes dont on nous parle aussi, on peut dire que le métier tel qu'on le connaît actuellement de livreur, taxi, chauffeur routier, peut disparaître ou être modifié.

Mais globalement, on ne sait comment réfléchir. Tout ce que l'on sait, c'est que selon une étude anglaise, sur les 140 dernières années, il semble qu'il y ait finalement eu plus d'emplois créés que la technologie n'en a détruits. Par exemple, lorsque la révolution industrielle, au 19^{ème} siècle, a introduit les machines textiles, la productivité horaire a augmenté de 50%, ce qui a abaissé le coût, provoqué une chute des prix, et ainsi augmenté la demande de textiles, qui a quadruplé entre 1830 et 1900, entraînant une hausse de l'emploi.

Une autre étude nous a paru intéressante. On la trouve dans la revue *Pouvoirs*, n°170, parue fin 2019. Le compte-rendu en est fait par Philippe Askenazy, directeur de recherches au CNRS (auteur de *Partager les richesses*), et Francis Bach, directeur de recherche à l'INRIA.

Les auteurs commencent par nous mettre en garde sur les discours alarmistes sur l'emploi émis par certains qui y ont en réalité un intérêt. Ainsi, des courants qui ont comme programme d'instaurer un revenu universel (un revenu qui serait distribué d'office à tout le monde, que l'on travaille ou pas) ont tout intérêt à nous alarmer sur les pertes d'emplois massives auxquelles il faudrait s'attendre avec le développement de l'intelligence artificielle. *“L'instauration d'un revenu universel, nous disent Askenazy et Bach, avait déjà été avancée dans les années 1960 comme remède à la disparition du travail générée par l'intelligence artificielle. Aujourd'hui, des coalitions hétéroclites convergent pour promouvoir cette fable sur les deux rives de l'Atlantique. Actifs en Europe depuis les années 1980, notamment sous l'impulsion de Philippe Van Parijs, certains tenants historiques du revenu universel y trouvent un nouvel argument pour vanter leur outil, censé être la version néolibérale de la protection sociale (...). En France, la vision d'un manque structurel d'emplois induits par l'IA et les robots est portée également à gauche par des (ex-)socialistes ou des écologistes”*.

Effectivement, on se souvient, au cours de la campagne présidentielle de 2017, du pré-candidat du PS, Benoît Hamon, qui avait axé son argumentation sur l'idée que les nouvelles technologies, parmi lesquelles l'intelligence artificielle, allaient supprimer massivement de l'emploi, ce à quoi il répondait en proposant l'instauration d'un revenu universel.

Dans cette version de gauche, expliquent Askenazy et Bach, *“la promesse est là d'une forme de libération de l'homme de la contrainte du travail ou de l'exploitation. Cette émancipation serait assurée par une taxe sur les robots qui aiderait à financer un revenu universel décent. Outre-Atlantique, on retrouve pêle-mêle des démocrates, des libertariens et des milliardaires de la Silicon Valley (Elon Musk, Mark Zuckerber - Facebook) pour défendre le couple IA-robots/revenu de base.”*

Les milliardaires des GAFAs y voient leur intérêt financier, puisqu'il faudra beaucoup de robots pour que le revenu universel soit plus conséquent. Quant aux gens "de gauche" à la Benoît Hamon, ce sera utile, selon Askenazy et Bach, pour : *“revivifier la protection sociale,*

remplacer les dispositifs sociaux ciblés et - écoutez bien, car là, c'est valable pour les milliardaires comme pour les aspirants représentant des pauvres - éviter une remise en cause de l'ordre capitaliste qu'ils dominent".

D'ores et déjà, l'IA a créé des emplois : l'apprentissage des machines, dans une première étape, on l'a vu et revu, a nécessité et nécessite encore de les nourrir par des entrées fastidieuses, où l'humain doit leur dire *"ça c'est un chat, ça ce n'est pas un chat"*. Travail qui se fait depuis chez soi, à coups de clics évidemment payés au lance pierre. L'IA a également considérablement augmenté la demande en énergie électrique. Mais notent nos auteurs, *"cette gourmandise en énergie préoccupante fait d'ailleurs partie des secrets les mieux gardés par les grands groupes industriels"*.

Le plus gros de l'IA déjà en place sert à cibler l'envoi de publicités, à trier les infos qu'on nous envoie. Mais cela a-t-il un impact sur l'emploi ? Pas évident. Pour que l'IA franchisse vraiment de nouveaux pas de manière massive, il lui faudra franchir plusieurs obstacles. Là où la santé (médecine), la sécurité (voiture autonome), entrent en jeu, il faudra qu'elle convainque, en plus des patients et des utilisateurs, les organismes certificateurs, qui donnent l'autorisation d'utiliser dans le public, mais aussi ceux qui décident des lois. Des accidents comme les récents crash du Boeing 737, dus à une nouvelle IA qui n'avait pas été assez clairement indiquée aux pilotes, pour économiser du temps d'apprentissage, peuvent faire réfléchir. La voiture autonome, pour exister réellement, devra nécessiter que les GAFAs et les constructeurs d'automobiles se partagent leurs connaissances, puisque les voitures devraient se piloter en échangeant directement entre elles. Mais les GAFAs sont-elles prêtes à prendre le risque que les constructeurs d'automobiles deviennent leurs concurrents ?

Enfin, Askenazy et Bach regardent la durée du travail à l'échelle de l'histoire. Et ils observent qu'après une longue période où ce temps de travail annuel avait reculé, cette réduction s'est arrêtée dans les pays avancés depuis le début de ce 21^{ème} siècle. Et même, sur une vie entière, l'âge de la retraite reculant, les années de vie au travail augmentent.

Le taux d'emploi, le pourcentage des gens au travail par rapport à ceux vivants, entre 15 et 64 ans, est au plus haut : il atteint des records dans les pays où la démographie est très faible (Japon, Allemagne) mais aussi là où elle est forte (Royaume-Uni). Même en France qui a un fort taux de chômage, le nombre d'emplois a fortement progressé, et le taux d'emploi a atteint 65,4% en 2018, au lieu de 60,7% en 1995, quand commence internet.

Dans la période à venir, estiment les auteurs, deux secteurs devraient demander de la main d'œuvre nouvelle : la dépendance, les vieux qui ne peuvent plus être autonomes, et l'adaptation au changement climatique. Mais ni pour l'un ni pour l'autre, on ne voit aujourd'hui possible un remplacement complet des humains par l'intelligence artificielle ou les robots. Dans l'état actuel des choses, ils pourront aider, mais pas remplacer. Enfin, les pays riches n'ont plus de gros réservoir de main d'œuvre : plus d'exode rural, c'est déjà fait ; la participation des femmes au travail, c'est pareil ; les immigrés, ce n'est plus de mode, et la démographie est en berne. On peut donc penser qu'on aura encore besoin pour pas mal de temps du travail humain, en tout cas dans le cadre de la société actuelle.

Askenazy et Bach concluent de tout cela qu'on ne peut rien dire de sûr sur ce qu'il en sera de l'avenir. Quand les distributeurs de billets de banque sont apparus dans les années 1970, expliquent-ils, on a annoncé une baisse des emplois dans les banques ; c'est le contraire qui s'est produit, car les coûts de sécurité ont baissé, une forte concurrence est apparue entre

banques et elles ont multiplié les agences. Ce n'est qu'un demi-siècle après, aujourd'hui, qu'une baisse de l'emploi s'amorce, avec les banques en ligne. Tout ce que l'on peut dire, c'est que l'IA n'implique pas par nature une disparition du travail.

Quelques conclusions

Pour nous qui sommes profondément socialistes, dans une société débarrassée du capitalisme, l'IA ne serait pas un problème mais une aide de plus pour alléger le travail humain, qui se doit d'être partagé entre tous. Une aide en particulier pour diminuer le travail contraint, libérant du temps pour pouvoir avoir une vie sociale, politique, culturelle.

Quant au soi-disant danger pour les humains d'être dépassés par une intelligence "forte", "*le cerveau ne fonctionne pas comme un ordinateur, et la mémoire n'est pas un disque dur*", nous dit le neuroscientifique Hervé Chneiweiss, président du comité d'éthique de l'Inserm. "*Je sais bien que des travaux existent pour fabriquer des ordinateurs qui miment le mode de fonctionnement du cerveau, par exemple le Human Brain Project. Malheureusement, les neurones vus par un informaticien sont aux vrais neurones ce qu'un Airbus est... au goéland : les deux volent, mais la comparaison s'arrête là ! Un Airbus ne pourra jamais se poser sur le flanc d'une falaise... Entre l'intelligence artificielle et l'intelligence naturelle, il y a le même gouffre.*

"La première fonctionne avec des algorithmes, c'est-à-dire une série de formules mathématiques. Et ceux-ci ne se reconfigurent jamais : une fois écrits, ils fonctionnent selon les instructions - et les biais - de leur concepteur. Alors que notre cerveau change d'avis tout le temps (...). Notre cerveau est bien plus puissant que n'importe quel ordinateur, on peut le transporter partout, il demande à peu près deux bananes par jour pour fonctionner sans être branché, et il donne plus de plaisir que n'importe quel ordinateur" (Télérama 02/10/2019)

Nous n'avons rien à craindre donc de la machine. Mais nous avons à craindre, par contre, de la société dans laquelle nous sommes, qui donne la richesse et le pouvoir à quelques-uns, qui se dissocient des autres et œuvrent à maintenir le fossé de l'inégalité et à l'aggraver. Une inégalité supplémentaire nous attend donc avec les nouvelles machines qui se profilent.

Le problème n'est donc pas vraiment nouveau pour nous, qui sommes conscients des limites et des dangers dus au fonctionnement social. Si danger il y a pour notre vie privée, mais aussi pour les libertés en particulier politiques, il est inhérent à cette société où une minorité détient le pouvoir, et pas aux techniques qui existent. Depuis que l'État existe, il a utilisé toutes les nouveautés techniques pour accroître et perfectionner son pouvoir.

Pour nous qui sommes de véritables socialistes, l'IA est un progrès qui devrait être mis au service de tous et de la vie collective. Dans une société libérée de l'exploitation économique, et débarrassée des appareils de méfiance et de surveillance d'État, l'IA sera un outil parmi d'autres et comme les autres, dont il n'y a rien à craindre, et que l'on regardera comme on regarde un vélo ou un téléviseur. Tout le monde sera conscient de ce dont il s'agit quand on use ou qu'on parle de l'IA, et plus personne ne sera impressionné par l'utilisation du mot "intelligence" dans l'expression "intelligence artificielle".

Sources

Nicolas Spatola, L'intelligence artificielle, L'Opportune 2018
Pouvoirs, L'intelligence artificielle, Seuil, n°170 2019
Papiers La revue de France Culture, Les révolutions de l'intelligence n°26 octobre-décembre 2018
Libération hors-série, L'IA au cœur de l'humain, décembre 2018 février 2019
La Vie hors-série, Les promesses de l'intelligence artificielle, septembre 2018
Laurence Devillers, Des robots et des hommes, Plon 2017
François Sureau, Sans la liberté, Tracts Gallimard 2019

Jean-Luc Schwartz, Google Duplex : l'illusion d'humanité, Pour la Science n° 102 septembre 2018
Jean-Paul Delahaye, IA : un apprentissage pas si profond, Pour la Science n° 488 juin 2018
Alison Gopnik, Quand les machines apprendront comme des enfants, Pour la Science n° 478 août 2017
Yoshua Bengio, La révolution de l'apprentissage profond, Pour la Science n° 485 juillet 2016
Frédéric Alexandre et Serge Tisseron, Quels sont les vrais dangers de l'IA ? Pour la Science h-s n°87 avril 2015

octobre 2019