

POURQUOI L'INTELLIGENCE ARTIFICIELLE A-T-ELLE BESOIN DE L'HUMAIN ?

Serge SOUDOPLATOFF et Pierre-Louis DESPREZ



IL ÉTAIT UNE FOIS DARTMOUTH...

L'Intelligence Artificielle est officiellement née en 1956, lors d'un séminaire qui réunissait à Dartmouth le gratin des chercheurs et penseurs en informatique, dont John McCarthy,

Marvin Minsky, Herbert Simon et Claude Shannon. Le thème du séminaire était « qu'est-ce que l'intelligence? ». Il s'agissait pour tous ces chercheurs de voir comment les paradigmes de l'informatique pouvaient être utilisés pour comprendre l'humain.

Il y avait une philosophie sous-jacente assez intéressante : utiliser un langage artificiel (celui de l'informatique) pour comprendre l'intelligence humaine, en évitant le paradoxe du dictionnaire qui a besoin de mots pour définir d'autres mots. Résultats : le terme d'Intelligence Artificielle a été créé, tous les langages de programmation, ainsi que les méthodes de modélisation des connaissances, en ont découlé.

Cette compréhension de l'intelligence que souhaitaient accomplir les participants du séminaire se retrouve encore de nos jours, par

exemple dans la bouche de Ishiguro, l'un des grands chercheurs sur les robots humanoïdes, professeur à l'Université d'Osaka : « Je construis des robots pour comprendre ce qu'est la conscience »¹. L'IA est déjà présente dans plein de moments de notre vie.

Ce qui distingue un *spam* d'un *email* acceptable est un moteur d'IA. Ce qui vous donne des recommandations lorsque vous effectuez des achats sur Internet est un moteur d'IA. Les compagnies d'assurance utilisent de plus en plus des moteurs d'IA pour estimer vos risques.

Antoine Compagnon, professeur de Littérature au Collège de France, explique aux chercheurs en littérature comment intégrer le *big data* dans la critique textuelle, en l'associant au « flair » de la recherche qui consiste à trouver un objet de recherche pertinent, l'IA faisant le reste.

L'intelligence artificielle pour le meilleur et pour le pire

Nous étions quelques chercheurs dans les années 1990 à travailler sur le thème de l'IA. Lorsqu'on nous en demandait la définition, nous avions coutume de répondre : « L'intelligence artificielle, c'est une intelligence débarrassée du risque de la bêtise humaine. » À cette époque, personne n'avait peur de l'IA ; sauf peut-être certains experts qui se demandaient si leur métier ne risquait pas de disparaître. Ce ne fut d'ailleurs pas le cas, les métiers qui ont disparu à cette époque concernaient surtout les tâches de manipulation de l'information avec peu de valeur ajoutée ; mais les experts, eux, sont restés.

En 2020, lorsqu'on nous pose la même question, la réponse est : « L'intelligence artificielle est ce qui plonge l'expertise dans une économie d'abondance ». Autrement dit, l'expertise n'est plus seulement dans la tête des experts, lesquels sont rares, mais devient disponible partout et à bas coût. Lorsque le *web* est arrivé, l'information est devenue abondante et à bas coût, puis avec le *web2.0*,

1 Cf. <https://www.youtube.com/watch?v=uD1CdjlrTBM>

c'est la connaissance qui est devenue accessible partout et à bas coût, Wikipedia en étant un des plus évidents exemples. Finalement, l'IA permet de réaliser le même mouvement avec l'expertise. Par exemple, *medoctor*² est un service en ligne gratuit basé sur une IA qui permet de faire un premier niveau de qualification de votre état de santé.

L'IA est la parfaite continuité de tout ce que l'informatique a créé depuis plus de soixante ans. Nous pourrions même situer les racines de l'informatique à l'invention de l'alphabet, lorsque des êtres humains ont inventé un système de symboles abstraits afin de décrire non seulement les objets du monde, mais aussi des nombres, des mathématiques, des concepts, etc.

Puis sont arrivés dans les années 1990 les « systèmes-experts ».

Il s'agissait d'une modélisation structurelle de l'expertise, avec des règles logiques de type « si - alors ». Pour modéliser la connaissance d'un expert, un nouveau métier a été inventé : cogniticien. Celui-ci s'asseyait à côté de l'expert, le regardait agir, lui posait des questions, et enfin rentrait tout ceci dans le modèle pour en dériver un programme informatique connu sous le terme de moteur de règles. Ceci posait deux problèmes à

l'expert : cela lui prenait du temps, et surtout il n'était pas forcément heureux de partager son savoir. Un expert aime bien vivre dans une économie de rareté. Par exemple, dans les années 1990, Renault ne vendait pas de voitures avec boîte automatique ; la raison était que les concessionnaires ne savaient pas réparer ces boîtes, et donc n'en faisaient aucune promotion. Un système-expert d'aide à la réparation a été développé. Aujourd'hui, les valises de diagnostic

Pour modéliser la connaissance d'un expert, un nouveau métier a été inventé : cogniticien. ”

2 Cf. <https://www.medoctor.com/>

automobile sont banalisées, et plus personne ne sait l'expertise qui y figure. L'IA est partout, mais banalisée.

En parallèle, d'autres chercheurs travaillaient sur des outils statistiques, avec une philosophie fondamentalement différente : il n'y a aucune connaissance du domaine d'application dans les algorithmes, seules les données sont utilisées. Jean-Paul Benzécri, le fondateur de l'analyse des données, le résume bien : « le modèle doit suivre les données, et pas l'inverse ».

À partir des modèles statistiques, une catégorie d'algorithmes innovants a vu le jour et a donné des résultats prometteurs : les réseaux de neurones. En s'inspirant de la manière dont les synapses s'interconnectent, cette classe d'algorithmes se base sur des graphes de relations qui portent des poids dont les valeurs sont calculées en fonction d'un jeu de données d'apprentissage. Ce sont ces algorithmes que l'on retrouve dans les systèmes de reconnaissance de formes : identification d'images, voitures autonomes, etc.

2011, la grande rupture: la convergence de trois innovations

Trois éléments totalement indépendants s'associent :

ImageNet : une première grande base de données d'images taguées.

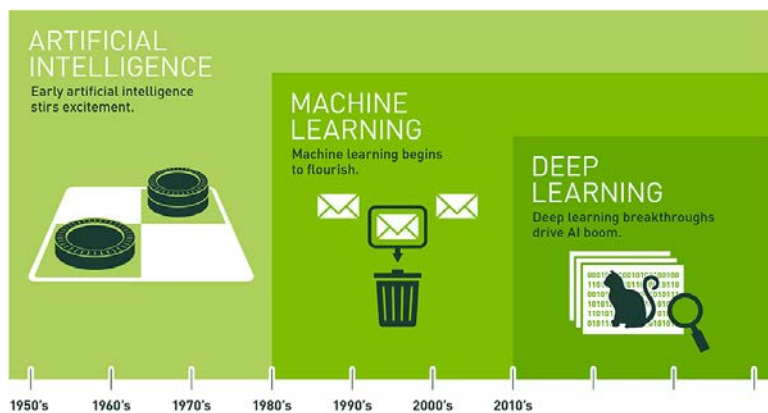
L'IA a besoin de données pour fonctionner. Un ordinateur sait reconnaître un chat ou un chien parce qu'on lui a montré des milliers de photos de chats ou de chiens, et qu'il a des algorithmes suffisamment sophistiqués pour être capable de dire sur une image s'il s'agit d'un chat ou d'un chien. Pour apprendre à la machine, il faut des données taguées. C'est ce qui s'est passé en 2011, avec ImageNet³, la première grande base de données d'images taguées. Les réseaux de neurones convolutifs : une nouvelle génération d'algorithmes statistiques.

3 Cf. <http://www.image-net.org/> aujourd'hui plusieurs millions d'images. Des concours d'algorithmes sur cette base sont organisés régulièrement.

Les algorithmes statistiques les plus puissants étaient les réseaux de neurones, dont l'architecture s'appuyait sur le fonctionnement des neurones dans le cerveau. Une autre classe d'algorithmes a vu le jour, les réseaux de neurones convolutifs, qui s'inspirent du fonctionnement du cortex cérébral. Ces algorithmes ont pu grandement améliorer la qualité des résultats obtenus jusqu'alors, entre autres dans des domaines dynamiques comme la reconnaissance de vidéos en temps réel, élément fondamental lorsqu'il s'agit, par exemple, de reconnaissance de personnes, d'intentions ou bien de conduite de véhicules.

Les GPU : une classe d'ordinateurs ultra rapides dotés d'une architecture en parallèle de type carte graphique utilisée pour les jeux vidéo. Ces algorithmes nécessitent une puissance de calcul énorme. Or, il se trouve que les cartes graphiques des ordinateurs sont bien plus puissantes que les processeurs eux-mêmes, tout comme le cerveau fonctionne d'autant mieux qu'il intègre des processus en parallèle. Un nouveau type de cartes graphiques est arrivé sur le marché, les GPU, qui se trouvent maintenant embarqués dans les systèmes d'IA.

Ces trois révolutions concomitantes ont donné un énorme coup d'accélérateur à l'IA, et ont ouvert la voie du *big data*, du *machine learning* et du *deep learning*. Ce sont trois disciplines imbriquées les unes dans les autres. Ce schéma, qui vient de la société NVIDIA qui est le leader en terme de GPU, l'illustre parfaitement : l'IA est ce qui permet à une machine de jouer au jeu de GO ; l'apprentissage est ce qui lui permet de distinguer entre un *spam* et un *email* acceptable ; et l'apprentissage profond est ce qui permet de détecter qu'une image contient un chat.



Since an early flush of optimism in the 1950s, smaller subsets of artificial intelligence – first machine learning, then deep learning, a subset of machine learning – have created ever larger disruptions.

L'IA, elle aussi, a besoin d'apprendre !

L'IA a besoin d'apprentissage pour fonctionner, donc de beaucoup de données. Plus les données sont nombreuses, plus le résultat sera fin. Il y a deux grandes catégories d'apprentissage : les apprentissages supervisés, où des données taguées sont rentrées dans la machine, et les apprentissages non supervisés, où la machine est censée apprendre toute seule sans avoir besoin d'un humain pour lui fournir des exemples. Ces apprentissages sont encore pour la plupart au stade de la R&D.

Qui possède ces données ? Les GAFA et les BATX, bien sûr. En une journée de 2020, les internautes publient 4,7 milliards de mises à jour sur *Facebook*, écrivent 5 milliards de tweets, font 5,7 milliards de requêtes sur *LinkedIn*, postent 70 millions de photos sur *Instagram*⁴. A l'inverse, l'état français qui devrait gérer l'ensemble des données médicales afin d'offrir de meilleurs services de prévention de santé, a du mal à créer un dossier médical personnalisé. Cela montre à quel niveau de pouvoir les GAFA sont arrivés et le chemin à parcourir par les administrations pour être au niveau.

4 Cf. <https://www.webfx.com/internet-real-time/>

L'IA pour les génies

Le célèbre AlphaGO de Google a réussi à battre le champion du monde de GO, dans sa version avec apprentissage. La génération suivante est partie de zéro : l'ordinateur a joué des milliards de parties contre lui-même, et en a déduit son algorithme par essais-erreurs. Du coup, cette version 2 a battu la version 1. Mais cet apprentissage non supervisé a réussi parce que les règles du jeu de GO sont très simples, et qu'un être humain les a codées dans la machine. Si l'on remplace ce problème par, par exemple, la prise de décision dans un comité de direction, voire dans un conseil des ministres, la complexité du problème rend un apprentissage non supervisé impossible aujourd'hui.

C'est donc grâce aux bases de données que les programmes d'IA deviennent de plus en plus performants. Les données sont les actifs structurants des sociétés d'aujourd'hui. Les deux grandes nations qui l'ont compris, Chine et USA, sont engagées dans une compétition sur la constitution de grandes bases de données et de compétences en IA. La Chine le fait, poussée par l'État Chinois qui a lancé un grand plan pour être le leader mondial en 2050, les USA par les *startups* et les GAFAs, sans doute en collaboration avec la *Defense Advanced Research Projects Agency* (DARPA) et d'autres administrations américaines.

L'IA pour tout le monde

Prenons un problème très représentatif du dilemme actuel en Intelligence Artificielle : la capacité d'une voiture autonome de prendre une décision importante dans un contexte de choix entre deux solutions dramatiques. Par exemple, comment choisir entre écraser trois piétons ou bien tuer les quatre passagers dans la voiture ?

Dans une IA de type structurelle, on pourrait imaginer des morceaux de code du genre « si ... alors... » Mais la complexité de toutes les possibilités rend une telle programmation impossible. Et puis, qui pourrait concevoir les règles ? Quel programmeur prendrait le risque d'écrire un tel code ?

Dans le cas statistique, peut-on vraiment concevoir un système basé sur des données d'apprentissage, et surtout comment capter

ces données ? Pour répondre à ce problème, des chercheurs du MIT ont créé un site web, *The moral Machine*⁵, qui propose des scénarios assez morbides à tout internaute, en lui demandant de faire un choix entre deux drames. Le site propose également à tout chercheur de définir ses propres scénarios. Une étude a ensuite été effectuée sur trois scénarios, comparant les réponses de 70 000

participants venant de 423 pays différents⁶. Les résultats montrent un lien entre faible mobilité relationnelle (je reste dans ma tribu) et le rejet du sacrifice pour le bien commun. Ceci illustre une des tendances de notre monde actuel, où la communauté est l'élément sociologique le plus structurant, bien plus importante que par exemple le département, voire la nation.

Des chercheurs du MIT ont créé un site web, « The moral Machine », qui propose des scénarios assez morbides à tout internaute, en lui demandant de faire un choix entre deux drames. ”

Pourquoi l'IA fait-elle peur ?

- L'IA est invisible, sauf au travers de représentations effrayantes comme les robots humanoïdes.
- La crainte d'une surveillance de la part de systèmes totalitaires, lesquels semblent fleurir en ce moment.
- La crainte humaine du changement, surtout induit par une technologie perçue comme une menace.
- La peur de voir « la machine » remplacer l'humain jusque dans ses fonctions cognitives, peur qui est nourrie par le mot « intelligence », pas forcément très bien choisi, même associé à artificielle...

5 Je conseille vraiment au lecteur d'aller « jouer » sur ce site, qui est proposé en dix langues, rien que pour sentir la difficulté que pose le design des systèmes d'IA.
<http://moralmachine.mit.edu/>

6 Cf. <https://www.pnas.org/content/117/5/2332>

Les sentiments à l'égard de l'IA sont ambivalents ; c'est un exemple supplémentaire du couple amour-haine. Les humains seront-ils capables de co-construire intelligemment le monde de demain avec l'IA ? Pour ce faire, il faut arrêter la peur qui nous empêche de nous transformer. La pire peur est celle de notre propre peur.

Des menaces réelles et fantasmées

- De nombreux métiers voient leur pratique bouleversée par l'arrivée de cette technologie, l'IA va détruire des emplois.
- La part des robots va augmenter, et sans doute occuper des espaces de décisions peu complexes.
- L'utilisation de l'IA par des gouvernements totalitaires induit de la surveillance de masse, de la création de fiches individuelles par corrélations de plusieurs traces laissées sur les réseaux sociaux.
- La dépression de l'humain face à l'appétit de connaissance : pourquoi étudier, pourquoi créer, pourquoi inventer, si les IA savent faire tout ça ?
- L'asymétrie entre les pays qui possèdent beaucoup de données et les autres, d'où risque de cyber-guerre, mais aussi de conflits de territoires liés aux migrations humaines.
- La nostalgie typiquement française d'une période mythologisée où l'humain nous semble avoir été au centre de tout et au-dessus de tout.

La disparition programmée de certains métiers

Les experts s'accordent tous pour prédire que la généralisation de l'IA va avoir un énorme impact sur l'emploi, et va créer beaucoup de dégâts. La crainte la plus commune est le remplacement de beaucoup de tâches par les robots. Les chauffeurs de taxi, les conducteurs de poids lourd, les manutentionnaires, tous sont condamnés au chômage. Des études futuristes prédisent quasiment l'apocalypse, comme celle de MacKinsey qui prétend que, vers 2030, 30% des salariés verront leurs rôles remplacés par des machines. Ces

études sont nombreuses, et un petit coup de Google permettra au lecteur de se délecter.

Pour essayer de se projeter dans le futur, prenons un exemple : le camion autonome. Une vidéo de Mercedes⁷ présentant son camion autonome montre qu'il y a toujours un conducteur dans la cabine ; seulement il ne touche pas au volant en permanence. Cette situation rappelle ce qui se passe dans la cabine d'un Airbus : à savoir un système qui est là pour assister le pilote dans sa charge de travail, et de manière différente selon les phases. Autrement dit, les chauffeurs routiers ne vont pas disparaître, ils deviendront des techniciens en charge de l'opération du véhicule. Leur principale compétence sera celle d'un opérateur : avoir une vision systémique

de la machine, être capable de prendre des décisions rapidement, et enclencher les bons niveaux d'automatisme selon les phases du trajet (en ville, sur autoroute, etc.). On voit déjà le problème classique : *quid* de ceux qui refusent de se transformer ; vont-ils se retrouver au chômage ? Le même problème s'est posé

Le seul domaine irrationnel où l'IA pourrait s'approcher de l'humain est le champ des émotions. ”

chez les pilotes d'avions. Au début des premiers Airbus, lorsque le manche à balais a été remplacé par un *joystick*, il y eut des voix pour dénoncer la perte du bon vieux temps où l'on pilotait les avions « avec les fesses », selon l'expression consacrée. Puis les pilotes sont devenus des ingénieurs systémiciens, et, du moins jusqu'à la crise de la Covid-19, le monde était plutôt en manque de pilotes.

Une autre raison de penser qu'il y aura toujours un conducteur dans le camion est le fait que les cinq niveaux d'autonomie des

7 Cf. https://www.youtube.com/watch?v=skrBb5g_zT4

véhicules⁸ ne sont pas encore tous opérationnels, en tout cas pas dans tous les contextes : il restera des contextes où le conducteur devra prendre la main. Lors du *Consumer Electronics Show* (CES) 2020 à Las Vegas, qui est maintenant le lieu où toutes les nouveautés en véhicule autonomes sont présentées, il était clair que la seule grande catégorie de véhicules professionnels qui proposait des véhicules autonomes opérationnels était les engins de chantier. Il est vrai qu'il s'agit d'espaces clos, où tous les niveaux d'autonomie peuvent être atteints.

Nous voyons se dessiner quelque chose de classique : ceux qui sauront se transformer avec la nouvelle technologie survivront. Les autres changeront de métier, et ceux sur le carreau devront avoir un revenu minimum, qui est une idée qui devient de plus en plus présente. Cette idée n'est pas neuve, le premier dans l'histoire connue à l'avoir mise en œuvre est Cyrus le Grand, vers 530 avant JC⁹.

Les robots vont-ils prendre le pouvoir ?

Une des grandes menaces vient du rôle que peut jouer l'IA lorsqu'elle est personnifiée dans des robots. L'arrivée de robots guerriers, déjà utilisés dans certains conflits, est un élément qui nourrit cette crainte.

Il faut distinguer l'IA faible de l'IA forte. L'IA faible est ce que nous voyons actuellement : systèmes de décision intelligents basés sur du *big data*, *machine learning*, qui permettent de réaliser voitures quasi autonomes, systèmes de reconnaissance d'images, etc.

En revanche, si l'on compare la technologie avec un être humain, nous sommes très loin devant la machine. Il faut des centaines de milliers d'images de vaches pour qu'une machine puisse apprendre ce qu'est une vache, alors qu'un enfant a besoin de beaucoup moins. Nous, les humains, avons développé quelque chose que la machine n'a pas encore : l'intuition, qui est notre extraordinaire capacité de comprendre, et d'avoir raison, quand nous avons peu de données, ce qu'une machine ne sait pas faire. Le seul domaine irrationnel où

8 Cf. https://fr.wikipedia.org/wiki/V%C3%A9hicule_autonome#Classification, sur les niveaux d'autonomie (5 ou 6 selon les pays)

9 Cf. https://fr.wikipedia.org/wiki/Cylindre_de_Cyrus

L'IA pourrait s'approcher de l'humain est le champ des émotions. Déjà, il existe des systèmes qui permettent de détecter des émotions uniquement en décodant le signal parole sur des paramètres tels que *pitch*, intensité, et pas du tout par la compréhension de ce qui se dit. Ces systèmes sont opérationnels ; par exemple, à Copenhague, les opérateurs téléphoniques de l'urgence sont assistés par un système qui détecte si la personne au bout du fil présente un risque de crise cardiaque¹⁰.

L'IA forte représente l'utopie qu'on construira des machines qui auront les mêmes capacités de raisonnement que les êtres humains. La série télévisée Suédoise *Real Human* illustre très bien ce que pourrait être une telle IA.

Tout laisse à penser que l'IA restera faible sur le long terme. L'intelligence d'un robot est actuellement basée sur des matériaux,

comme le silicium, c'est-à-dire des molécules symétriques. Comme l'a montré Pasteur, la vie est basée sur des molécules asymétriques, et les biologistes savent bien que trouver le modèle 3D d'une ARN n'est pas un exercice simple. En terme de complexité, la machine est loin de l'être humain et, pourrait-on dire, se situe sur une branche parallèle. Si d'un côté les processeurs

Si maintenant l'ordinateur peut battre le champion du monde de GO, il n'est pas encore au niveau d'aller fêter ça avec ses copains ordinateurs au bar du coin. ”

tournent plus vite que notre cerveau (plusieurs GHz alors que, si l'on prend comme base un centième de seconde pour que l'information passe les synapses, notre cerveau fonctionne à 100 Hz, soit un rapport de 1 à 40 millions), d'un autre côté la reproduction non seulement de la structure interne du cerveau, mais de toutes

¹⁰ Cf. <https://www.businessinsider.com/danish-startup-developed-technology-that-predicts-heart-failure-2018-5>

les complexités biologiques liées aux échanges entre molécules complexes en trois dimensions, n'est pas encore au niveau des machines les plus performantes. Ajoutons à cela que toutes les tentatives d'en terminer avec la logique binaire des ordinateurs, comme la logique ternaire, ou la logique floue, ont échoué ; le 0 – 1 est vraiment résistant. La seule grande voie vraiment disruptive à l'heure actuelle est l'arrivée des ordinateurs quantiques. Mais c'est un sujet naissant, très pointu, qui fait dire beaucoup de bêtises ; bien malin celui qui déjà peut expliquer les bases de la mécanique quantique de manière simple¹¹.

Comme il est joliment dit, si maintenant l'ordinateur peut battre le champion du monde de GO, il n'est pas encore au niveau d'aller fêter ça avec ses copains ordinateurs au bar du coin. Et puis, la machine n'est pas capable de trouver toute seule son alimentation. Le magnifique film de Kubrik, *2001 l'odyssée de l'espace*, est encore d'actualité : l'humain peut toujours débrancher l'ordinateur.

Les états asserviront-ils les citoyens grâce à l'IA ?

En décembre 2015, le journal *The Guardian* révèle que la société anglaise *Cambridge Analytica* a, via ses algorithmes, influencé les primaires présidentielles du Parti Républicain en faveur de Ted Cruz, à sa demande. Les algorithmes ne devaient pas encore être parfaits, car ce fut Donald Trump qui fut choisi par les électeurs Républicains. Ils furent améliorés, et utilisés pour influencer à la fois l'élection de Donald Trump, et le Brexit. Grâce à un lanceur d'alerte, le 17 mars 2018, *The New York Times* et *The Observer* publient un même article¹² révélant que, grâce à une application qui a aspiré des millions de données d'utilisateurs *Facebook*, *Cambridge Analytica*¹³ a pu profiler ces utilisateurs et proposer de leur envoyer des

11 Rien que l'intrication quantique, qui est une des bases de la physique quantique, ne peut se comprendre qu'en possédant une connaissance mathématique poussée des espaces de Hilbert. La définition exacte est « Deux particules sont dites dans un état intriqué lorsque l'état des deux particules n'est pas factorisable en un produit vectoriel de deux états à une particule ». Cf. https://fr.wikipedia.org/wiki/Intrication_quantique

12 Cf. <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>

13 Le patron de *Cambridge Analytica* à cette époque n'était autre que Steve Bannon.

publicités ciblées sur certains thèmes politiques biaisés, influençant ainsi l'élection en faveur de Trump, ou bien du Brexit.

Les menaces sont donc présentes, et ont déjà eu lieu. Ceci ne peut qu'augmenter la tentation des états d'utiliser l'IA pour influencer la géopolitique mondiale.

Chacun peut en avoir un avant-goût, grâce à un évènement qui s'est produit chez *Facebook* le 18 mars 2020 : plusieurs personnes ont vu leur message censuré, alors qu'ils ne contenaient rien de vraiment répréhensible. Beaucoup étaient de simples liens vers des articles de journaux pas moins inoffensifs que d'habitude. *Facebook* a dû s'excuser et expliquer que, à cause du confinement en cours¹⁴, la censure avait été plus automatisée que d'habitude, par manque de moyens humains¹⁵. Deux choses intéressantes sont à déduire : la première est que *Facebook* ne semble pas prêt pour le télétravail ; la deuxième, que malgré la forte compétence de *Facebook* en IA, les opérateurs humains font encore mieux que l'IA, au moins sur ce sujet.

Pour nous protéger contre la surveillance des états, peut-être devons-nous tous utiliser des astuces, par exemple pour que les caméras de détection ne puissent pas nous identifier, comme ces activistes qui se peignent le visage¹⁶ ? Pourquoi pas, surtout que les techniques ne manquent pas¹⁷ et que même des *designers* se sont mis à inventer de tels camouflages¹⁸ !

14 En ce mois de mars 2020, la Bay area autour de San Francisco a été confinée avant la Californie, alors même que la Maison Blanche ne se résout toujours pas à confiner tout le territoire des États-Unis. C'est une preuve de plus que le système politique américain est plus proche du local que du global.

15 Cf. <https://www.lesechos.fr/tech-medias/hightech/facebook-google-et-twitter-au-defi-de-la-moderation-sans-supervision-1186514>

16 Cf. <https://www.theguardian.com/world/2020/feb/01/privacy-campaigners-dazzle-camouflage-met-police-surveillance>

17 Cf. <https://www.businessinsider.com/clothes-accessories-that-outsmart-facial-recognition-tech-2019-10>

18 Cf. <https://cvdazzle.com/>

Des opportunités à ne pas rater

L'IA aide à gérer la complexité d'un monde global

Si l'on regarde des grands défis qui se posent à la planète, comme la pauvreté, le climat, l'environnement, l'éducation, nous trouvons toujours un point commun : le monde est passé en un siècle d'un peu plus d'un milliard d'individus à plus de sept milliards.

Sept milliards d'astronautes dans le même vaisseau spatial terre, cela signifie un accroissement énorme des interactions¹⁹. Nous côtoyons plus de gens dans les rues, sur les routes, dans les trains, les avions, etc. Gérer la complexité croissante d'un tel système ne peut plus se faire sans le numérique.

En 1965, l'aéroport d'Orly comptait 6 millions de passagers ; en 2019, 31 millions, auxquels il faut ajouter les 79 millions de Roissy-Charles de Gaulle qui n'existait pas en 1965, nous avons un trafic multiplié par 18. Rien que la gestion correspondante des avions, depuis l'approche jusqu'au sol, débarquement et embarquement, ne peut pas se faire sans le numérique.

Dans un monde complexe, il faut atteindre l'homéostasie pour résister. Le corps humain est un parfait exemple d'atteinte de cette homéostasie ; et c'est parce que les éléments de notre corps s'échangent en permanence de l'information que nous sommes toujours en vie malgré les agressions du monde extérieur, et même que nous sommes capables de croître, au moins dans les premières années de notre vie. Ceci est

aussi vrai dans un système même non biologique : pour atteindre l'homéostasie, il faut que l'information, et aussi la connaissance, puissent circuler librement. Le numérique, surtout depuis Internet,

Il faut éliminer tout ce qui bloque la fluidité des trois éléments : information, connaissance, expertise. ”

¹⁹ Rappelons que le nombre d'interactions entre n individus est $n * (n-1) / 2$ donc de l'ordre du carré du nombre d'individus. Imaginer le carré de 7,5 milliards est vertigineux, même divisé par deux.

est ce qui nous a permis de faire face à cette complexité croissante. L'augmentation probable du nombre de personnes dans le monde, donc, au carré, du nombre d'interactions, va nécessiter plus que des systèmes de gestion de processus : cela va nécessiter de l'expertise. Et c'est là où l'IA devient fondamentale, dans sa capacité à apporter de l'expertise, partout, à tout le monde et à tout moment. Ceci nécessite néanmoins des changements d'organisation : il faut éliminer tout ce qui bloque la fluidité des trois éléments : information, connaissance, expertise. Les modèles de gouvernance en silo sont une catastrophe. La hiérarchie n'est pas du tout adaptée. Et surtout, mal français, les modèles basés sur la méfiance sont à éliminer. La bonne utilisation de l'IA nécessite un changement profond de nos structures et modes de pensée.

L'IA au service de la santé

Prenons un exemple dans la santé : les déserts médicaux. La solution proposée actuellement, à savoir favoriser les consultations à distance, est typique de l'attitude qui consiste à raisonner comme par le passé, mais avec les outils modernes, ce qui au pire est contre-productif, au mieux inutile, mais ne fonctionne jamais.

Ce qui peut être gagné en terme de temps de trajet par le médecin ne peut pas compenser le manque de contact physique, voire une présence parfois nécessaire pour faire un diagnostic.

Pourquoi ne pas imaginer alors une formation de nouvelles classes de médecins à une vision synthétique du corps humain, sans avoir besoin de connaître par cœur le nom de tous les os, à qui seraient fournies deux choses : des capteurs à bas prix connectés à des *smartphones*, qui enverraient les données à un système d'IA à distance, lequel retournerait au médecin le ou les quelques diagnostics, avec un système de dialogue guidé lui permettant si nécessaire d'affiner son propre diagnostic ?

Tous les éléments existent. Ainsi, lorsqu'il s'agit d'analyse d'images, l'IA se révèle plus performante que les médecins dans certains do-

maines²⁰. Les capteurs miniaturisés commencent à couvrir tous les domaines, qu'il s'agisse de scanners échographiques portables, de tensiomètres, de suivi du rythme cardiaque, etc. On commence même à voir des séquenceurs d'ADN à moins de 1 000\$. Si tous les éléments existent, le plus dur sera néanmoins de convaincre le corps médical de la pertinence d'une telle révolution. Les temps des décisions humaines, surtout lorsqu'elles sont collectives, n'est vraiment pas le temps de la technologie.

Des logiciels et services en *open source* pour tous

Ce qui rend l'IA intéressante, c'est que beaucoup de chercheurs sont dans la logique ouverte du *net*, et que beaucoup de logiciels d'IA sont donc disponibles en *open source*. Citons par exemple :

- *Tensorflow*, un système *open source* développé initialement par *Google Brain*, extrêmement complet²¹.
- *IBM Watson*, développé par IBM, lequel a analysé des millions de document dans divers domaines²².
- *Torch*, un *framework open source* centré *machine learning* maintenu par des ingénieurs de *Facebook*, *Twitter* et *Google*²³.
- *Scikit-Learn*, un ensemble de logiciels orienté *data mining*²⁴.

Cette liste n'est en rien exhaustive, et il suffit de chercher sur Google pour en trouver d'autres. De plus, beaucoup d'entreprises, voire de start-up, proposent des interfaces de programmation (API) qui permettent d'avoir des services sans avoir besoin de coder. Par exemple, Google permet de renvoyer des valeurs émotionnelles à un texte qui lui est fourni *via* une API.

.....
 20 Des concours sont régulièrement organisés entre médecins et IA sur l'analyse d'images. Cf. <https://technologiemedica.net/2018/07/04/ia-bat-les-medecins-dans-un-concours-en-neuro-imagerie/>

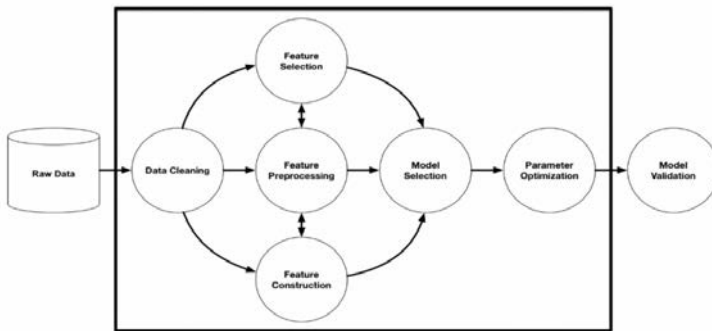
21 Cf. <https://www.tensorflow.org/>

22 Cf. <https://www.ibm.com/watson>

23 Cf. <http://torch.ch/>

24 Cf. <https://scikit-learn.org/stable/>

Néanmoins, avoir un logiciel ou des API ne suffit pas à résoudre le problème. Les données d'apprentissage sont fondamentales, mais il faut bien faire attention à leur qualité. L'IA est comme la cuisine : on ne peut pas faire de bons plats avec des ingrédients médiocres. Les données ne sont pas tout, il faut également travailler sur ce que l'on veut en extraire. En d'autres termes, pas d'IA sans objectif. Le graphique ci-dessous illustre le chemin à parcourir lorsqu'on veut développer un système d'IA, bien avant même de le faire fonctionner :



Enfin, ne rêvons pas d'une IA qui résoudra tous les problèmes, cela est irréaliste. En revanche, on assistera de plus en plus à des IA locales, qui résoudront des problèmes dans des contextes bien définis et pour lesquelles les données d'apprentissage seront spécifiques. Les entreprises s'orientent maintenant vers ce modèle : pas de grande IA générique, mais des IA locales et spécifiques.

L'IA est attendue dans les domaines de l'éducation et de l'environnement

Dans le monde de l'éducation, l'IA peut également apporter de l'enseignement là où les enseignants sont rares. Mais l'IA peut faire mieux : *via* des questions-réponses, une IA peut faire un profil des connaissances de chaque étudiant et orienter les futures interactions en leur apportant ce qui leur manque, en se basant sur une

modélisation de ce qu'est un espace de savoir²⁵. La Chine est d'ailleurs très volontariste dans ce domaine.

L'IA, de par sa capacité à traiter des masses de données, est capable de faire des prédictions climatiques beaucoup plus fines que les modèles actuels. ”

Dans le domaine de l'environnement, l'IA, de par sa capacité à traiter des masses de données, est capable de faire des prédictions climatiques beaucoup plus fines que les modèles actuels. Par exemple, la société *CarbonTracker* effectue un suivi mondial des pollutions industrielles, en se basant sur des images satellitaires²⁶.

Quelle éthique humaine pour l'intelligence non-humaine ?

L'éthique nous place dans le domaine de la pratique et non pas des grandes valeurs morales, tels le bien, l'honnêteté, le courage, le respect etc. Elle est pragmatique. C'est ce dont nous avons besoin pour frayer avec l'intelligence non-humaine dénommée Artificielle.

Trois principes directeurs

Le premier principe est celui de la réciprocité. Un acteur économique développant des IA, pour agir éthiquement, devrait se poser la question non pas du bénéfice de l'utilisateur, lequel débouche finalement sur un gain financier, mais celle du bénéfice pour l'humanité. Penser aux autres, et non pas à un autre qui utilisera et paiera directement ou indirectement. Laissons à l'économie ses mécanismes, éclairons-les avec des questions qui la dépassent et qui l'englobent, d'autant que, sait-on jamais, les humains ont toujours la capacité de ne plus acheter de voitures fussent-elles autonomes,

25 Cf. <https://www.aleks.com/>

26 Cf. <https://carbontracker.org/carbon-tracker-to-measure-worlds-power-plant-emissions-from-space-with-support-from-google-org/>

de ne plus utiliser *Tinder* pour trouver leur moitié, de préférer la main du chirurgien aux scalpels de Da Vinci. La réciprocité rejoint l'universalité. C'est un combat, rien n'est gagné d'avance. C'est cela, être humain.

Le deuxième principe qui en découle est celui de l'empathie.

Si l'IA que je développe met au chômage 5 millions de conducteurs de camions aux États-Unis, il est toujours possible de chercher des solutions pour eux. À terme le coût de leur employabilité et de leur reconversion est toujours un sujet pour la collectivité. San Francisco fait rêver, les clochards sur les trottoirs beaucoup moins. Par chance l'empathie n'est pas une question de morale, elle est liée aux neurones-miroirs découverts par le biologiste Rizzolatti. L'humanisme, c'est-à-dire poser l'Humain au centre de tout comme ultime critère, est une relation de réciprocité.

Le troisième principe éthique c'est que l'usage fait toujours la différence.

Et il y a un bon et un mauvais usage de l'IA. Le sujet, ce n'est pas l'IA, mais l'usage que nous en faisons. Le sujet, c'est l'humain, comme depuis toujours. Remplacer un contrôleur de train par un *chatbot* est parfois plus efficace, mais pas toujours. En voyageant un jour dans le Shinkansen, je me rendis compte que j'avais oublié mes baguettes sur le comptoir du *sushi-shop* qui se trouvait dans la gare. Le contrôleur s'en rendit compte lorsqu'il me vit commencer à attraper les petits rouleaux à deux doigts. D'un geste il me signifia d'attendre. Une minute plus tard, il m'apportait deux baguettes. L'occurrence étant très faible et les probabilités encore plus – sauf avec des Occidentaux tête-en-l'air qui préfèrent le système D aux règles de l'art –, il est évident que le *chatbot* eût été moins efficace. Mais l'IA peut apprendre, elle le prouve chaque jour.

Du bon et du mauvais usage

Qu'est-ce donc qu'un mauvais usage sur le plan éthique et non pas sur celui de la *form factor* ni de l'ergonomie *user friendly* ?

Il y a mauvais usage lorsque l'Autre disparaît de nos consciences. Le Progrès – avec majuscule s'il vous plaît ! – a eu ses heures de noblesse en Europe lorsque les découvertes, les inventions et les

innovations étaient pensées, réalisées, adaptées pour le plus grand nombre. Jules Ferry nuisit beaucoup aux paysans qui utilisaient leur marmaille pour semer, sarcler, moissonner, vendanger et nourrir les cochons. Il envoya leurs rejetons à l'école par la force légale, et certains devinrent des personnalités éminentes. La fierté les paya de retour, enfin... pas dans tous les cas. Le principe d'universalité aurait-il disparu ? Voire ! En Inde l'innovation est dénommée inclusive et non pas exclusive. La communauté est consubstantielle à l'acte individuel. Nous avons beaucoup à réapprendre des « pauvres ». L'ennemi de l'éthique, c'est l'individualisme, évidemment, c'est-à-dire le fait d'ériger la Personne seule et unique au centre de tout.

Que l'IA contribue à détecter une tumeur mieux que les meilleurs P.H.P.U. ne fait pas débat. Chacun y trouve son intérêt. Il nous en faut plus ! Nous avons besoin de la technique, inventée par les

humains, pour résoudre ce que l'humain ne parvient pas à trouver.

Ce sont les deux infinis pascaliens : si finis et si infinis sommes-nous ! Espérons donc que les virologues et les technologues vont coopérer pour prédire les prochaines pandémies virales. Car ce ne sont pas les cris dans le désert de l'OMS depuis l'épidémie de SRAS, puis l'épisode du H1N1 qui

ont réussi à transformer les budgets annuels des nations dites « développées » au bénéfice de la prévention sanitaire. Espérons que les développeurs informatiques d'algorithmes vont coopérer avec des philosophes et des éthiciens pour choisir qui, du vieillard ou de l'adolescent, la voiture autonome emballée doit écraser. Humains, trop humains, et si faibles. Mais humains, très humains, pour aborder les pires questions, prendre les plus insupportables décisions.



**L'ennemi de l'éthique,
c'est l'individualisme. ”**

Que serait encore un mauvais usage de l'IA ? Son rejet. Rien n'est pire que l'ignorance au pouvoir, la plus grande démocratie au monde nous le prouve chaque jour à propos de la gestion du confinement dû à la Covid-19. Le rejet de l'IA, c'est le rejet de

l'humain qui a le talent de l'inventer. Quelle tyrannie pourrait lui mettre un frein ?

Et si on éduquait ? Le mot fait souvent sourire ceux qui oublient qu'ils sont le produit d'une éducation. Abdiquer sur l'humanisme, c'est ne plus croire en l'humain. Nous avons oublié le Progrès ? Remettons-le au centre, non pas dans quelque instance lointaine qui finit toujours par tourner en rond, mais dans le labo, sur la paillasse, lors de l'accompagnement d'une thèse de doctorat, lors d'une levée de *love money*. L'IA s'emballe et fait peur ? Utilisons notre plasticité cérébrale pour imaginer les scénarios du pire et du meilleur, puis marchons en crabe pour décider. À quoi cela sert-il d'avoir 100 milliards de neurones si ce n'est pour trouver là où la difficulté est extrême ? Nous ne sommes pas des animaux aux comportements appris et sur-appris, il nous reste peut-être les 25% de conscience auxquels Spinoza pensait lorsqu'il écrivait : « Nous sommes des cynocéphales dans les trois quarts de nos actions. » On a parfois gagné avec moins de 25% d'ingéniosité.

Assigner des finalités à l'IA

Et si on finalisait l'IA sur les problèmes que, même dans notre double infini, nous avons quand même du mal à résoudre ? Par exemple peut-on stopper le réchauffement climatique autrement que par une régression ou des taxes qui stimuleront la ruse des coupables et les rendent encore plus efficaces grâce à leurs 100 milliards de neurones ? En quoi l'IA pourrait-elle nous y aider, et au plus vite, avant que le maïs ne crame en avril chaque année ? En quoi l'IA pourrait-elle nous aider à accélérer le développement du vaccin ? Impossible ? Tant mieux ! Que l'IA s'y colle, il y a même des milliards à la clé, et cela fera un beau sujet de thèse pour les doctorants.

Et si l'on n'oubliait plus d'enseigner les quelques principes éthiques empiriques fondamentaux qui tiennent sur les doigts d'une seule main, au lieu d'en faire des commissions, des chartes imbuvables, et des manifestes à la noix ? On s'attache trop aux questions de la liberté individuelle (cf. *Stopcovid*) où le débat éthique est trop souvent ravalé au niveau de la protection des données indivi-

duelles, qu'il ne s'agit pas d'éluder. Ce faisant, on oublie le débat éthique sur les fondamentaux de notre humanité : la réciprocité, la dignité, l'intégrité de l'autre. Et ne pas décider pour lui s'il ne le veut pas. Et ne pas forcer subrepticement les peuples et les pays à laisser des IA décider pour eux. Le résultat d'une décision pertinente ne compte pas plus que le process qui a permis de l'élaborer. Cela s'appelle l'apprentissage à vivre ensemble.

Le débat sur l'IA n'est pas qu'économique ni technique, il est politique, car il transforme profondément la société, et non pas à la marge. ”

Mais si on cessait de croire que l'IA c'est forcément un humain augmenté ? Cédric Villani, dans son rapport sur l'IA et la France, répète la *doxa* économique sans aucun recul. En guise de démonstration, on a droit à une suite d'injonctions au nom de la compétition entre les États. Je m'attendais à mieux de la part d'un scientifique couronné qui a pratiqué le doute heuristique, l'in-

validation de l'hypothèse, la falsifiabilité d'un énoncé. En clivant ainsi le débat entre pour et contre, et en prenant position comme si l'autre n'existait pas, comme s'il fallait convaincre à tout prix, il fait l'impasse sur un débat, sur l'expression de la contradiction, et donc sur l'émergence par la dialectique de la tierce solution. Attitude qui ne vaut pas mieux que les comités d'éthique créés par les grands groupes internationaux, lesquels se font juge et parti. Le débat sur l'IA n'est pas qu'économique ni technique, il est politique, car il transforme profondément la société, et non pas à la marge. Il concerne le bien public, qui n'appartient à personne, mais nous concerne tous. Il y aura sans doute peu de citoyens qui s'opposeront à l'apport de l'IA dans le diagnostic médical. Il y en a davantage qui seront récalcitrants quand ils y verront l'équation IA = destruction de mon gagne-pain. Le débat est posé, en situation, et non pas en vérité.

Le débat est-il si complexe ? Quelle bonne nouvelle ! Il va falloir apprendre à penser du nouveau, à trouver des raisonnements adaptés, à éviter la bêtise et l'idéologie. Il va falloir faire preuve d'humanité, et non d'humanisme. Il va falloir parler d'ici et maintenant, anticiper des futurs possibles, au lieu de convoquer un Humanisme de la Renaissance comme critère de jugement de ce qui est bon et de ce qui ne l'est pas, de ce qui est adapté à notre humanité et de ce qui la menace.

Vers la réciprocité ! Un regard éthique nous oblige, pour finir, à parler en réciprocité. Si l'humain, comme il l'a toujours fait, peut s'augmenter des capacités de la technique, la technique, inversement, doit être conçue avec des finalités humaines, en donnant une place à l'humain. La vision d'une compétition est délétère. La vision d'une complémentarité de bénéfices est heuristique. Humain augmenté ou IA humanisée ? Les deux !

« Aucun savoir ne se forme sans un système de communication, d'enregistrement, d'accumulation, de déplacement qui est en lui-même une forme de pouvoir [...]. Aucun pouvoir, en revanche, ne s'exerce sans l'extraction, l'appropriation, la distribution ou la retenue d'un savoir. À ce niveau, il n'y a pas la connaissance d'un côté, et la société de l'autre, ou la science et l'État, mais les formes fondamentales du "pouvoir-savoir"* . »

Michel Foucault

« "Écoutez vos sentiments" recommandait l'humanisme. "Écoutez les algorithmes ! Ils reconnaissent vos sentiments !" recommande le dataïsme . »**

Yuval Noah Harari

* Confer Foucault, 2015, p. 77.

** Massachusetts Institute of Technology.