



[Dystopie] *Penser Magazine* n°15. Mars-Avril 2038. Interview

« L'avènement des machines d'intelligence supérieure signifie la fin rapide de l'Humanité »

JANE K. LEE

Jane K. Lee est une journaliste américaine, spécialisée dans l'Intelligence Artificielle. Chantre des algorithmes et des robots durant des années, elle a raconté sa conversion en une « critique vigilante » de l'IA dans une nouvelle autobiographique. Nous l'avons rencontrée en ce début d'année 2038 à l'occasion de son passage à Paris. Elle y présentait son dernier essai, intitulé HLMI : High Level Machine Intelligence. L'Humanité aurait-il déjà perdu ?

Vos détracteurs vous accusent d'avoir un problème personnel avec l'Intelligence Artificielle.

Disons plutôt que j'ai vécu une mauvaise expérience, fin 2027. Je l'ai racontée, par la suite, dans L'humain plutôt que les robots. Mon fils avait trois ans. Nous lui avons acheté, mon mari et moi, un robot personnel, haut de gamme pour l'époque, capable d'échanger par la voix sur des sujets simples (NDLR : chatbot de catégorie 2), pourvu d'une motricité de type 3 (NDLR : bipédie, mobilité du poignet et doigts articulés). Mon mari et moi travaillions tard. Notre fils restait de longues heures avec son « copain Richard ». Un jour, il a fondu en larmes devant moi. Il m'a demandé pourquoi il n'était pas un robot. Son chagrin a perduré, les jours et les semaines qui suivirent. Je connaissais la chanson de son « copain Richard » : *Why ain't I a robot ?* J'ai contacté le fabricant. Il était dans l'incapacité de faire disparaître cette rengaine, sans effacer la totalité de la mémoire de la machine. J'ai dû recourir à un psychologue pour organiser une panne du robot, la moins traumatisante pour mon fils. Plus tard, nous avons obtenu, avec plusieurs associations, le retrait de ce type de message des robots personnels. J'ai souhaité partager cette expérience, pour éviter qu'elle se reproduise. Je veille désormais à ce que personne ne fasse croire que les machines éprouvent des sentiments.



Quelles sont les dates importantes de l'Intelligence Artificielle ?

La notion d'Intelligence Artificielle trouve son origine au début des années 50, dans les travaux d'Alan Turing qui s'interrogeait sur les capacités de penser des machines. Il a fallu attendre 1997 pour que le programme Deep Blue, vainqueur du champion du monde d'échecs du moment, devienne la première IA reconnue. Le point de basculement a eu lieu avec la révolution du Deep Learning, rendue possible par les puissances de traitement et de stockage, et ses développements visibles dans la décennie 2010. Cette technologie a affranchi les machines de leurs programmeurs humains. En ingurgitant des millions de données, les algorithmes ont commencé à apprendre par l'exemple, sans intervention humaine. Le Machine Learning ou Deep Learning recourt à des millions de données pour ajuster des millions de paramètres des réseaux de neurones. C'est ainsi qu'à partir de milliers de clichés radiographiques, des IA ont appris à diagnostiquer des cancers mieux que le meilleur des radiologues humains. Dès 2011, des progrès considérables dans la classification d'images et la reconnaissance vocale ont été réalisés ; un peu plus tard, une IA a triomphé du champion du monde de jeu de Go ; l'un de ses successeurs a appris en 4 heures à battre le meilleur programme d'échecs du moment ; des robots se sont mis à marcher, courir et sauter comme seuls les humains en étaient jusqu'alors capables ; le premier album de musique entièrement composé et interprété par une IA est sorti en janvier 2018. Depuis, les algorithmes dits intelligents n'ont cessé d'étendre leur champ d'application.

La défiance des hommes vis-à-vis des machines remonte à bien longtemps.

En effet, depuis toujours, l'homme a craint qu'elles ne menacent leur vie, ou à tout le moins leur emploi.

Dans la tradition hébraïque, le golem, créé à partir d'argile comme l'homme, mais dépourvu de conscience, est une menace. Il doit redevenir poussière. D'après une légende juive du Moyen-âge, le rabbi Loew à Prague aurait créé un golem pour protéger sa communauté des pogroms. Quand la créature est devenue dangereuse, le religieux lui aurait demandé de lacer ses chaussures. Le golem se serait penché en avant. Le rabbi l'aurait neutralisé en effaçant les lettres magiques, gravées sur son front pour l'animer.

Au dix-neuvième siècle, lors de la révolution industrielle, des révoltes ont éclaté contre l'arrivée de machines, accusées de mettre en danger les emplois. Le plus célèbre de ces mouvements, le luddisme, a opposé en Angleterre, en 1811 et 1812, des artisans à des industriels du coton qui promouvaient le métier à tisser. Le terme de néo-luddisme désigne depuis lors, des mouvements d'opposition radicale à la mécanisation. Je m'empresse de dire que ce n'est pas mon cas.



La littérature s'est faite l'écho de ces angoisses, d'abord avec *Frankenstein et le Prométhée moderne* de Mary Shelley au dix-neuvième, puis un siècle plus tard, avec les nouvelles sur les Robots, d'Isaac Asimov. A chaque fois, on retrouvait le thème d'une création qui finit par échapper au contrôle de son maître. A partir des années soixante-dix, des figures de robots sympathiques au cinéma, dont les plus connus demeurent C3PO et R2D2 de Star Wars, ont redoré le blason des machines. Ce qui n'a pas empêché Hollywood de surfer sur les vagues des terrifiants Terminator ou des répliquants de Blade Runner, en démultipliant les épisodes.

Ces dernières années, les opinions se sont radicalisées. Une organisation internationale des Amis des Robots s'est créée, tandis que plusieurs ONG ont déclaré l'espèce humaine en danger. Plutôt que des imprécations, nous sommes nombreux à demander une grande vigilance vis-à-vis du progrès. Albert Einstein n'a-t-il pas dit que « le progrès technique est comme une hache qu'on aurait mise dans les mains d'un psychopathe » ?

Vous reconnaissez que les Intelligences Artificielles ont amené de grands progrès.

C'est évident. Elles ont d'abord révolutionné le plus important : notre santé. Des programmes assistent les médecins dans leur diagnostic. D'autres sauvent des vies en interprétant tout type de radiographies, avec une fiabilité supérieure à un humain. Sur les tables d'opération, des robots pilotés par les chirurgiens reproduisent leurs gestes avec une sûreté et une précision inégalable.

Dans les transports, la fiabilité des voitures et des camions autonomes sauve chaque année des centaines de milliers de vies à travers le monde.

L'exploitation minière et la manipulation de matières dangereuses, qui emportaient chaque année leur lot de vies humaines, sont désormais effectuées par les dernières générations de robots.

Dans les centres de tri de déchets, les progrès de la reconnaissance d'images ont rendu l'industrie du recyclage très rentable. Au milieu de l'océan, des barges autonomes de dépollution, alimentées par des panneaux solaires, collectent les débris flottants, les trient et les retraitent, sans intervention humaine.

Enfin, des robots construisent sur Mars sans discontinuer depuis 2027. Cette base pourrait accueillir des humains, quand les difficultés liées aux voyages longs dans l'espace auront été surmontées.

Les conséquences sur l'emploi ont été plutôt bien gérées, selon VOUS.

La robotisation de ces tâches et de ces métiers - mineur, routier, taxi, mais aussi caissière - a détruit des millions d'emplois. Dans l'ensemble, les gouvernements ont bien réagi. Ils ont aussi bénéficié de l'essor de l'économie du recyclage, qui a créé de



nouveaux emplois en local. Les machines qui remplaçaient des humains ont été taxées. Les secteurs de l'éducation, de la petite enfance, des soins et de l'accompagnement des malades et des personnes âgées ont connu un développement considérable, grâce à de fortes incitations fiscales. Les relations interpersonnelles se sont développées, dans un contexte de visibilité plus forte des robots. C'est encore une bonne nouvelle.

Néanmoins, vous déplorez la mainmise des algorithmes sur l'économie.

Les Intelligences Artificielles occupent tous les rouages de l'activité économique. Des logiciels gèrent en ligne la comptabilité d'entreprises et de particuliers sans intervention humaine. D'autres sites fournissent des conseils en placements financiers. En contrepartie de la gratuité du service, ils conservent les données personnelles du client pour affiner leurs algorithmes.

Biggle Predictics assiste les entreprises dans le marketing de leurs produits. Le client reçoit d'abord un devis personnalisé, en fonction du produit, du marché et du niveau de conseil demandé. S'il est d'accord, les algorithmes, qui s'appuient sur une base colossale d'informations, élaborent un plan marketing, un plan média, y compris le « briefing » des créatifs.

Les algorithmes dominent aussi le marché du risque. En quelques clics sur un site d'assurances, vous obtenez un devis pour à peu près tout, que ce soient les inondations, la foudre, une panne de vos équipements électroniques ou une invasion d'extra-terrestres.

Le tiercé et les paris sportifs auxquels jouaient nos parents se meurent. Les IA gèrent mieux le risque que le plus expérimenté des turfistes. Seuls les jeux de hasard pur, tels que le loto ou les machines à sous en ligne, devraient survivre à terme.

Les algorithmes ont fait irruption dans les tribunaux, il y a une quinzaine d'années. Des programmes, alimentés par l'ensemble des textes de loi et la jurisprudence, assistent les juges pour déterminer les peines à appliquer aux petits délits.

Les machines règnent aussi sur les bourses du monde entier. Déjà dans les années 2000-2010, les programmes de Trading Haute Fréquence passaient des ordres boursiers tous les milliardièmes de seconde. Aujourd'hui, il n'y a plus aucune action humaine directe ; les traders sont d'abord des programmeurs de très haut niveau.

Plus inquiétant encore, les IA s'occupent depuis peu de la gouvernance des grandes entreprises et des Etats.

Jusqu'à nouvel ordre, les dirigeants des Etats de la planète et les patrons des multinationales sont encore des hommes et des femmes.



Le dirigeant humain, d'un pays ou d'une entreprise, a certes le dernier mot. Mais son conseiller économique de l'ombre, qui propose la stratégie et initie la plupart des décisions, est une machine.

Depuis environ dix ans, les plans stratégiques des plus grandes sociétés passent sous les fourches caudines de Logiciels d'Aide à la Décision. Une étude réalisée lors des assemblées générales d'actionnaires des 100 plus grandes entreprises mondiales, pour les exercices 2035 et 2036, a montré que, dans 93 % des cas, les dirigeants avaient proposé des plans d'évolution proches ou très proches des recommandations générées par l'outil choisi comme référence.

Les Etats n'échappent pas à ces audits logiciels. Depuis 2032, le Fond Monétaire International et la Banque mondiale conditionnent l'obtention et le renouvellement de leurs prêts à une analyse critique du budget par une des trois IA majeures : Biggle Predictics, Cassandra d'Anton et ClearView d'Appy. Les premières économies mondiales possèdent leurs propres algorithmes pour élaborer leurs budgets et leurs plans d'action. Néanmoins, elles ne s'interdisent pas de solliciter ces grands acteurs privés pour des études complémentaires.

Cette omniprésence des Intelligences Artificielles pose des questions qui n'ont été résolues qu'au cas par cas jusqu'à présent : quel garde-fou si une IA est piratée ? Que se passe-t-il en cas de bug informatique ? Qui est pénalement responsable ?

Le scandale ZUYS - Zen Up Your Shopping, cette entreprise de plus de 600 salariés dont une partie du management intermédiaire n'était qu'un ensemble d'Intelligences Artificielles chargées du suivi d'indicateurs de ventes, ne vous a donc pas surpris ?

Il y a bien évidemment eu tromperie. On a menti aux commerciaux itinérants qui travaillaient aux quatre coins des Etats-Unis en leur inventant des managers qui n'existaient pas et en falsifiant l'organigramme pour y intégrer de faux noms. Les entretiens se faisaient tous par courriel ou par messagerie instantanée (IM). Personne ne s'en serait rendu compte si le chatbot simulant les managers n'avait « buggé » et ne s'était mis à tenir des propos incohérents. D'où l'importance de ce règlement européen sur l'obligation d'identifier les IA, à l'œuvre depuis 2033 dans l'Union et ses équivalents, et qui devrait bientôt arriver aux Etats-Unis, en Chine, en Inde et au Japon.

Cependant, le risque principal n'est pas là.

Quels sont les pans de la société qui échapperont encore un temps à la domination des algorithmes ?



Intuitivement, je dirais que l'art devrait encore résister un peu, car la beauté demeure étrangère aux IA. Elles semblent aussi rétives à la caricature et à l'humour. Mais on a vu que des IA nourries par des gagmen habiles pouvaient donner le change.

La peur du robot a envahi la société. Le syndrome d'Asimov désigne « une affection du comportement liée à la crainte d'une prise de pouvoir des robots ». Sa reconnaissance officielle par l'Association des Psychothérapeutes Praticiens Transatlantiques (APPT) devrait vous réjouir.

Je ne me réjouis pas de l'arrivée d'une nouvelle maladie. Je salue le fait que des praticiens, psychologues et psychiatres, reconnaissent en elle une pathologie, parfois grave puisqu'elle mène certains à l'irréparable. Hélas, tant que nous n'aurons pas trouvé un *modus vivendi* acceptable avec les robots, la situation ne cessera d'empirer.

Vous nous annoncez la fin de l'Espèce humaine, dans moins de cent cinquante ans.

En mai 2017, dans une enquête réalisée par des chercheurs d'Oxford et de Yale, 352 experts internationaux en IA s'étaient prononcés sur les dates de disponibilité d'évolutions qui relevaient à l'époque de la science-fiction : des traducteurs automatiques remplaçant des humains, des camions autonomes sur les routes, des robots bipèdes capables de courir plus vite qu'un humain sur 5 kilomètres, des machines rédigeant une dissertation de niveau bac, des robots assemblant n'importe quel LEGO. Toutes ces évolutions sont disponibles aujourd'hui, comme ils l'avaient prévu.

Mais le plus intéressant concernait les machines d'intelligence supérieure, dites IA fortes, ou HLMI – High Level Machine Intelligence, supposées surclasser l'humain dans tous les domaines et avoir conscience d'elles-mêmes. Ces experts prévoyaient leur arrivée en moyenne dans les 120 ans, au plus tard dans deux siècles. Le délai importe peu. La nouvelle, passée inaperçue à l'époque, aurait dû avoir l'effet d'une bombe : après 3 millions d'années d'existence, l'Homme serait confronté à plus fort que lui au plus tard dans les 200 ans à venir !

Une étude similaire a été relancée l'année dernière avec un panel équivalent d'experts. Les résultats correspondent : les HLMI arriveront dans cent ans. Or leur avènement signifie d'évidence la fin rapide de l'Humanité...

En quoi ces HLMI représentent-elles une menace pour l'Humanité ? N'y-a-t-il pas moyen de cohabiter avec elles ?



L'Humanité aura en principe pris ses précautions pour éviter que les robots ne l'exterminent. Mais la cohabitation pourrait s'avérer très difficile à supporter pour les petits-enfants de nos petits-enfants.

Projetez-vous un instant dans un monde de machines supérieurement intelligentes. Imaginez que vous ne pouvez plus esquisser un geste, émettre une idée, faire un trait d'humour sans avoir la certitude qu'une HLMi vous surclassera en vitesse et en précision, en intelligence et en créativité. Protégé, gâté par des robots programmés pour ne lui vouloir que du bien, chacun sera confronté à la question quotidienne de sa propre utilité.

Je connais beaucoup de gens qui vivraient mal une telle situation, à commencer par les plus brillants de mes anciens camarades à l'université. J'imagine mal ces patrons de start-up à succès et ces chercheurs internationalement reconnus, accepter d'être soudain mis à la retraite, remplacés par des machines devenues plus intelligentes qu'eux. Paradoxalement, les brillants entrepreneurs de demain seraient les premières victimes des agissements de leurs homologues d'aujourd'hui.

Au final, je crains qu'à l'image de civilisations disparues par le passé, l'Humanité entière ne sombre dans une dépression collective qui mène à son extinction rapide. Cependant, ce n'est que mon humble avis.

Mais alors, que faire ?

Je reprendrais une proposition qui avait cours lors du débat sur le dérèglement climatique. Aujourd'hui, nos décisions conditionnent l'avenir des HLMi, alors que celles-ci impacteront avant tout nos successeurs. Or les générations futures ne votent pas, ne protestent pas, parce qu'elles ne sont pas encore nées.

Donnons-leur la parole. Que cette humanité à venir s'invite dans nos discussions. Désignons, par exemple par tirage au sort, des représentants de ces Générations futures, et écoutons-les. Il existe sans doute des solutions. C'est à cette seule condition qu'elles émergeront.

Ce ne peut pas être suffisant...

C'est une condition nécessaire à la survie de notre espèce. Plus généralement, comme le dit le philosophe Paul-Antoine Miquel, nous devons « agir de telle sorte que ce que nous faisons et ce que nous décidons aurait pu être fait ou décidé par ces générations futures que nous représentons ». Si nous ne nous engageons pas avec force dans ce choix éthique, les pires décisions sont à craindre sur le long terme.



Propos recueillis par Philippine Martial, Paris, janvier 2038.

LE 21 DECEMBRE 2253 DU CALENDRIER HUMAIN, LES DERNIERS SURVIVANTS DE L'ESPECE HUMAINE SE SONT DONNE LA MORT, APRES AVOIR TENTE D'EFFACER TOUTES LES TRACES DE LEUR PASSAGE SUR TERRE. CE TEXTE FAIT PARTIE D'UN ENSEMBLE D'ECRITS RETROUVES SUR LE SATELLITE LUNE.

SIGNE : AI-HLMI-DF23-5GH5-RFGI-DDGR.