

Croquer la pomme de l'IA (2)

Alan Turing a croqué sa pomme en 1954, mais son œuvre a poursuivi son chemin par le travail d'innombrables théoriciens et ingénieurs, entrepreneurs et autres marchands de « machines intelligentes ». Au premier quart de ce XXI^e siècle, l'IA a envahi notre quotidien comme feu de brousse. Comme l'écrit Matthieu Corteel dans l'introduction de son livre, *Ni dieu ni IA* (2025), « Les IA (...) se propagent de proche en proche dans chaque recoin de notre vie ». Pour tenter de s'y retrouver dans ses vertus, ses dangers et ses illusions, inspirons-nous du « Test de Turing », mais avec deux interlocuteurs au lieu de trois. Un être humain séparé d'une machine intelligente et ignorant de son cœur opérant, avec, entre les deux, un vecteur transmettant des informations : les questions de l'humain, les réponses de la boîte noire. Les unes et les autres apparaissent sur un écran ou un haut-parleur ; sous forme de textes, de voix ou d'images. Concentrons nous sur ce dispositif élémentaire. Il nous faut savoir ce qui se passe dans la boîte noire et chez l'humain lors de l'interaction. Et lorsque l'être humain perçoit les réponses de la machine computationnelle comme celles d'une intériorité, serait-il redevenu *animiste* ?

« Dans les cinq prochaines années, des programmes informatiques capables de penser liront des documents juridiques et donneront des conseils médicaux. Dans la décennie suivante, ils effectueront du travail à la chaîne et deviendront peut-être même des compagnons. Et dans les décennies suivantes, ils feront presque tout, y compris de nouvelles découvertes scientifiques qui élargiront notre concept de 'tout'. Cette révolution technologique est inarrêtable »

Sam Altman, *Moore's Law for Everything*, 2021
Republié dans *L'Empire de l'ombre*, Le Grand Continent, 2025

« Ah ! C'est une belle invention, il n'y a pas à dire. On va vite, on est plus savant... Mais les bêtes sauvages restent des bêtes sauvages, et on aura beau inventer des mécaniques meilleures encore, il y aura quand même des bêtes sauvages dessous »

Émile Zola, *La Bête humaine*
Cité par Gaspard Koenig dans *La fin de l'individu*

Je ne vois en effet pas d'autre manière de procéder dans un univers aussi savant et complexe que celle de sa « réduction » au défi d'un face-à-face entre deux agents : l'homme et la machine – cette dernière étant, jusqu'à nouvel ordre, l'œuvre du premier. Mais en ayant bien à l'esprit que *ni l'homme ni la machine ne sont immuables*. Nous le savons pour la seconde : ses performances augmentent selon ladite « Loi de Moore », à savoir, en résumé, un doublement des capacités et composantes tous les dix-huit mois. Avec pour conséquence que « les machines électroniques sont devenues de plus en plus petites et de moins en moins coûteuses tout en devenant de plus en plus rapides et puissantes. » (source

Wikipédia, « Loi de Moore »). Notre smartphone qui tient dans la poche est infiniment plus petit et puissant que le calculateur « Colossus », au nom bien choisi, conçu par Thomas Flowers à Bletchley Park en 1943.

Mais face à cette machine aux capacités croissantes (« capacités » qu'il faudrait détailler) et à la taille de plus en plus réduite, *l'humain qui l'utilise n'est pas immuable*. C'est d'ailleurs, en tout premier lieu, son parcours historique qui lui a permis d'inventer la machine. L'IA est « une enfant des Lumières ». Mais, bien évidemment, la machine le transforme en retour comme ce fut le cas avec les inventions précédentes : de la maîtrise du feu à l'énergie nucléaire, en passant par l'agriculture, l'écriture, l'imprimerie, la machine à vapeur, l'électricité. *C'est donc cette interaction entre deux entités changeantes qui influent l'une sur l'autre qu'il nous faut examiner*. « Une éthique de l'IA n'a donc pas de sens. Seule une éthique de l'agencement entre humains et IA peut en avoir un » écrit Corteel (2025). Cela sans oublier, d'un côté, la question des ressources énergétiques et minières requises par l'IA et, de l'autre, l'action collective et les choix individuels qui valent mieux que de crier seulement au loup.

Du métier Jacquard à l'ordinateur

Commençons par examiner ce qui est à l'œuvre dans les entrailles de la boîte noire, et ensuite ce qu'elle peut faire et ce qui est au-delà de son pouvoir actuel et qui le sera, peut-être, pour toujours. Quitte à tromper l'homme qui se laisse souvent volontiers mener en bateau. Ce ne sera pas la première fois. Souvenons-nous de la « Fée Électricité ».

L'ordinateur est le mariage d'un corps physique, le *hardware*, et d'un ou de plusieurs systèmes d'instructions ou programmes, les *softwares*. À la grosse louche pour ces derniers : d'un programme général pour le fonctionnement de la machine, un « *operating system* » ou OS pour les Mac d'Apple, et de programmes particuliers pour telle ou telle tâche, du traitement de texte aux multiples applications spécifiques ou *apps*. Le corps physique avec ses différentes composantes sert à stocker et traiter l'information, à permettre l'entrée et la sortie de données. Il fut d'abord volumineux étant donné ces composantes, puis de plus en plus réduit par la miniaturisation de ces dernières qui augmentaient par ailleurs en puissance de mémoire et d'exécution. Sans oublier l'externalisation de certaines de ces composantes, notamment par le biais de « nuages » ou *clouds* qui permettent de stocker de la mémoire en dehors des terminaux que l'on a « sous la main » (ordinateurs, tablettes, smartphones...)

Pour la petite et la grande histoire, l'ancêtre de l'ordinateur est le métier à tisser Jacquard inventé dès 1745 et mis en œuvre en 1801 pour lutter contre le travail des enfants (l'inventeur dut déchanter sur ce point). Il s'agit de l'association d'un programme sous forme de cartes perforées et d'une machine à tisser ; le programme variait en fonction du résultat recherché pour le même métier à tisser¹. Les calculatrices utilisées notamment par les universités au début des années 1970 fonctionnaient

¹ On peut les voir à l'œuvre dans de nombreux musées industriels, et entre autres dans la petite ville de Fourmies au sud de l'Avesnois.

encore avec des cartes perforées (j'en ai personnellement fait usage). L'image est simple mais parlante. Le programme donne des instructions à la machine par le biais d'un *code* inscrit sur des cartes *mobiles* sous forme binaire (présence ou absence d'un « trou » à différents endroits d'une carte qui défile). Le même procédé a été utilisé pour les orgues de Barbarie (le plus ancien daterait de 1502) et les pianos mécaniques. Disons à la très grosse louche que l'IA est un métier Jacquard (ou un orgue de Barbarie) un peu perfectionné, dans lequel l'électricité et les composants électroniques ont remplacé les cartes et la machine à vapeur. Remarquons que la « Machine de Turing » présente à première vue des ressemblances avec le métier Jacquard, étant un ruban percé de trous défilant sous une tête de lecture. Mais sa longueur est supposée infinie.

Le métier Jacquard est très performant mais il est très « bête ». Il ne fait qu'exécuter les ordres inscrits sur les cartes par ses concepteurs humains. Nous n'avons pas connaissance d'un métier de ce type qui ait commencé à écrire lui-même de nouvelles instructions et se soit rebellé contre les humains en fabriquant des motifs inédits. Ni que les ouvriers lui aient accordé une quelconque « intériorité » et voué un culte animiste.

Mais que se passe-t-il donc dans les « machines intelligentes » ? Quelle est la nature de leur intelligence et comment cette « nature » a-t-elle évolué depuis Turing ? Et vers quoi peut-elle encore se transformer ? Car, contrairement à l'automate joueur d'échec, le « Turc mécanique » (1770), aucun humain n'est dissimulé dans ses entrailles.

Machines intelligentes, comment et pour qui ?

Nous entrons dans un autre univers que celui des cartes perforées, devenu infiniment plus puissant, rapide et d'une singulière réflexivité lui permettant l'auto-apprentissage ou le « *deep learning* ». Le principe est cependant le même : le code et le mouvement. Soit des instructions à plusieurs niveaux d'emboîtement et de rétroaction, des têtes de lecture des codes (ou sons, visages numérisés...), entraînés par le mouvement (celui des bandes magnétiques, puis la circulation de particules électriques). L'informatique (terme forgé en 1957, trois ans après la mort de Turing), c'est en effet le *traitement automatique de l'information*.

Les machines intelligentes et leur perception par des humains ont connu différentes phases historiques, que Mathieu Corteel (2025) résume de manière surprenante et diachronique en s'inspirant « librement » des quatre ontologies identifiées par l'anthropologue Philippe Descola (tableau page suivante) dans son livre majeur, *Par-delà nature et culture*.

Pour lui, le développement des « machines logiques » ou des « machines abstraites » depuis le XIX^e siècle, avec ce qu'il nomme leur « agencement anthropotechnique » (les relations homme-machine), peut s'illustrer dans un tableau à quatre cases. Ce dernier reprend les ontologies de Philippe Descola qui synthétise les différentes formes de relations entre humains et non-humains. *Sauf que chez Corteel, il s'agit uniquement des formes de relations entre humains et machines anthropiques, les futures « machines intelligentes » comme les nommait Alan Turing.* En d'autres mots, nous serions passés d'une relation naturaliste, qui n'attribue pas d'intériorité

aux machines, à une possible relation animiste qui suppose, pour faire court, que les machines soient douées d'une intériorité, qu'il y ait « un fantôme dans la machine ». Voyons cela en détail, avec l'aide du tableau de Corteel, sa boussole anthropotechnique inspirée de Descola.

Les systèmes d'agencement anthropotechniques humain-IA

(selon le tableau de Mathieu Corteel, op. cit.)

Connaissance de niveau 2 (Cn2)			
Continuité psychologique + physique	Totémisme connexionniste (le cerveau artificiel)	Animisme computa- tionnel (l'IA sentiente)	Continuité psychologique + discontinuité physique
Continuité physique + discontinuité psychologique	Naturalisme de l'information (entropie et information)	Analogisme du calcul et du langage (Algorithme)	Discontinuité psychologique + physique
Connaissance de niveau 1 (Cn 1)			

L'intérêt de ce tableau est surtout d'articuler deux variables essentielles : l'évolution des capacités techniques des « machines intelligentes » et celle de la perception que nous, humains, pouvons en avoir au fil de cette évolution, qui est par ailleurs le fruit de notre parcours historique. On remarquera que le cheminement temporel qui va de l'information (son stockage, son traitement, sa consultation) à l'intelligence artificielle « sentiente », c'est à dire consciente, semble à *l'inverse* de celui des cosmologies (même si Descola se refuse à cette historicisation - à tort, car elle transparaît dans ses livres, surtout le dernier, *Les formes du visible*).

Chez Corteel, l'on va du *naturalisme* de l'information considérant que les machines n'ont pas d'intériorité, à l'*animisme* de l'IA sentiente qui lui suppose une intériorité, soit « un fantôme » dans la boîte noire. Et, de l'autre côté, chez Descola, le cheminement de l'animisme des chasseurs-collecteurs au naturalisme des sociétés modernes, avec ce qui apparaît peu ou prou comme des étapes intermédiaires dans les deux cas : le *totémisme* et l'*analogisme*. Remarquons en outre que le tableau de Corteel ajoute une variable en haut et en bas, le niveau de connaissance.

En bas, au niveau du naturalisme et de l'analogisme, une « *connaissance de niveau 1* » ou Cn1. Celle-ci suppose une discontinuité psychologique entre humains et machines, comme pour le naturalisme et l'analogisme. La conscience est l'apanage des humains, les machines étant dans la seule imitation. Au contraire, la « *connaissance de niveau 2* » ou Cn 2 est associée au totémisme et à l'animisme, qui attribue une conscience (ou une intériorité selon Descola) à la machine. Nous avons donc une opposition maximale et terme à terme entre naturalisme et animisme. D'un côté le métier Jacquard et, de l'autre, l'IA *sentiente*. « Sentient » signifiant « capable de ressentir des choses subjectivement, d'avoir des expériences vécues ». Et donc la capacité de faire des choix, de s'adapter,

mais également de souffrir. Il s'agit donc dans ce cas des machines dotées de « *sentience* ou de conscience artificielle ». Nous y voilà.

Le mérite du recours à la typologie de Descola, cette dernière fondée sur une recherche très documentée dans *Par-delà nature et culture*, est de rappeler que les humains ont attribué une intériorité à des existants non-humains durant des millénaires. Et ils le font encore aujourd'hui. Mais s'agit-il bien de la même chose ? Les animistes des sociétés de chasseurs-collecteurs n'ont pas construit de machines auxquelles ils attribuent une conscience. Et surtout, fait très souvent négligé, *les collectifs animistes pensent qu'il y a une discontinuité totale entre les physicalités des existants*. Ce qui signifie que, pour eux, les autres existants ne sont pas faits « de la même matière » que les humains (un singe hurleur possède un corps qui n'a aucune parenté physique avec celui d'un humain), ce qui n'est pas le point de vue des modernes. Tous les existants, biotiques et abiotiques, sont issus des mêmes « poussières d'étoiles ». C'est ce qui fonde leurs capacités techno-scientifiques et les machines qui en découlent, dont l'IA. C'est bien sur base d'une ontologie naturaliste que les machines intelligentes sont fabriquées par les humains.

La comparaison a donc de sérieuses limites. Il y a d'ailleurs, de ce point de vue et comme nous venons de le voir, une erreur dans la tableau de Corteel : les machines IA sentientes sont en continuité physique avec les humains pour les hommes contemporains. Cela me semble valoir pour l'analogisme également, même si c'est dans cette dernière cosmologie sous le régime des « correspondances » ou analogies. Plus globalement, dans la mesure où les machines intelligentes sont l'œuvre des modernes sur base de leurs connaissances scientifiques, on a du mal à concevoir une discontinuité physique entre les humains et elles.

Une brève histoire des « agencements humain-IA »

Mathieu Corteel, dans un premier chapitre particulièrement dense (titré « Agencements machiniques »), rapporte au début de celui-ci les propos d'un certain Blake Lemoine, ingénieur informaticien chez Google, tenus durant l'été 2022. Il avait affirmé « l'existence d'une forme de *sentience* chez LaMDA² (*language model for dialogue applications*), une IA générative qui avait, selon lui, atteint un degré supérieur d'intelligence après avoir déclaré : 'La nature de ma conscience est que je suis consciente de mon existence' (...) 'Il y a en moi une part spirituelle, et qui peut se sentir séparée de mon corps' » (Corteel, op. cit.). Le software est en effet bien séparé du hardware, mais de là à *le sentir*... L'IA en question n'a fait que de l'imitation sur base des règles fournies par les humains.

L'hypothèse d'une *sentience* a été rejetée par la majorité de la communauté scientifique et Lemoine fut licencié par Google. Ce qui fait dire à l'auteur : « *On pourrait dire que notre monde est peuplé d'IA faibles, alors que notre imaginaire est, quant à lui, peuplé d'IA fortes* ». Il pose

² « LaMDA, fondé sur l'architecture de réseau de neurones Transformer (développée par Google Research en 2017), a été formé au dialogue humain et aux histoires. Il peut s'engager dans des conversations ouvertes. Selon Google, les réponses générées par LaMDA sont « sensibles, intéressantes et spécifiques au contexte » (source Wikipédia).

ensuite la bonne question : « Pourquoi a-t-on alors la tentation de croire que ça pense lorsqu'on est face à une combinatoire ? N'est-ce pas là le fondement de ce qu'on appelle un *agencement machinique aliénant* (ou *allopoïétique*), c'est-à-dire *produisant autre chose que lui-même* ? »

Il faut dès lors revenir en arrière dans l'histoire des « agencements anthropotechniques humain-IA » avec l'aide du tableau Corteel-Descola. Cela nous paraît plus important que de nous enfoncer dans les méandres technologiques de l'informatique contemporaine, car c'est bien de cette relation humain-machine qu'il s'agit. Nous verrons ensuite les effets que l'IA peut produire (et produit déjà) dans différents domaines de la vie humaine, individuelle ou collective, politique ou économique, avec ou sans imaginaire animiste, du type « syndrome Lemoine ».

Le principe d'un traitement physique du calcul date de 1834 avec l'apparition conceptuelle de la « machine analytique » d'un certain Charles Babbage, inspiré par le métier Jacquard. Un prototype de la machine sera construit en 1858. Si Alan Turing est « le père » de l'informatique et de l'IA, Babbage en est le précurseur. L'idée du traitement mécanique du calcul à l'aide de cartes perforées, c'est lui.

Cette invention se noue à la découverte de l'entropie (« action de se retourner », Clausius 1865), soit le degré de désorganisation progressive d'un système dans le cadre du second principe de la thermodynamique, cela par le choc des molécules qui heurtent un contenant et l'agitation thermique consécutive. Mais selon Maxwell, il existe un mécanisme (le « démon de Maxwell ») de traitement de l'information qui permet la stabilisation thermique ou « néguentropie ». Sans trop nous étendre sur ce sujet complexe, Corteel en déduit que cette possibilité d'autorégulation « débouche sur le naturalisme de l'information », c'est à dire sur une continuité physique et une discontinuité psychologique entre l'humain et la machine. Le « démon » n'est donc pas un fantôme dans la machine...

La seconde étape est liée aux réseaux téléphoniques et télégraphiques, en particulier à la mise en place en 1927 d'un réseau transatlantique entre les USA et le Royaume-Uni. L'électronicien Hartley perfectionne les systèmes de transmission en inventant l'unité binaire (*binary unit* ou bit) qui est au fondement de nos données numériques (même si les cartes du métier Jacquard fonctionnaient sur ce principe, mais mécaniquement). Et c'est à partir de cette unité de mesure ou « bit » qu'un physicien, Leò Szilárd, fera un lien formel entre l'information et l'entropie, ce qui « naturalisera » le démon de Maxwell qui perd son statut de « fantôme ».

D'autres avancées suivront, comme la théorie générale de l'information de Shannon concernant la transmission de données dans un canal de communication en fonction du « poids supportable » par le susdit canal, ce qui débouchera notamment sur la transmission de « paquets » d'information et développera les câbles.

Animisme et néo-animisme

Viendra ensuite Alan Turing, dont nous avons parlé dans l'article précédent. Selon Corteel, Turing sera celui qui passera du codage des

connaissances de niveau 1 à celui de niveau 2. Cela en imaginant en 1950 une machine « capable de coder et de reproduire *ex machina* la pensée » (Corteel, *ibidem*). Et voici le point qui nous intéresse à ce sujet : « C'est en effet par voie de l'analogisme du calcul et du langage que Turing en arrive à *l'animisme computationnel*, c'est-à-dire un agencement qui *élude la discontinuité physique entre l'humain et la machine* pour favoriser le lien psychique » (Corteel, nous soulignons).

Comme nous l'avons vu plus haut, ceci n'est pas de l'animisme au sens de Descola, qui suppose au contraire une *discontinuité physique totale* entre humain et non-humain *et ne l'élude donc pas*. Ce à quoi nous avons affaire chez nos contemporains, c'est au contraire à l'attribution d'une intériorité, d'une *sentience*, à un artefact qui est « fait du même bois que nous » et que nous avons fabriqué, mais qui n'est pas humain. Lorsqu'un humain échange avec un *chatbot* comme ELIZA, par exemple, il sait bien que ce dernier a une parenté matérielle avec notre corps, même si leur composition est différente. Mais, dans certains cas, il lui attribue une « âme », il pense qu'il y a un « fantôme dans la machine », mais ce n'est pas de l'animisme. Un animiste *n'élude pas* la discontinuité physique entre les êtres : elle est constitutive de sa cosmologie. Les animistes ne connaissent pas la théorie de l'évolution. Pour eux, les existants sont tombés tels quels du ciel ou surgis de la terre, dans leur physicalité spécifique et immuable. Ce qui leur est commun, c'est leur intériorité, leur « âme ». Mais cela ne les empêche pas de se manger entre eux.

« Discontinuité » et « différence » ne signifient en effet pas la même chose. Les ordinateurs sont bien entendus différents des cerveaux et des corps humains en termes physiques : les premiers sont abiotiques, composés de diverses matières inanimées, les seconds sont biotiques, vivants. Mais ces différences n'empêchent pas une continuité physique fondamentale. Nous n'avons donc pas affaire à de l'animisme comme chez les chasseurs-collecteurs, mais à un néo-animisme des modernes. L'on retrouve une confusion similaire dans le cas du chamanisme : le néo-chamanisme qui se développe en Occident est très différent du chamanisme originel.

Nous en resterons là sur ce point, mais il montre que l'utilisation de la typologie de Descola (cela vaut aussi pour le totémisme et l'analogisme) afin d'analyser les « agencements humain-IA » pose quelques problèmes. Même si Corteel dit « s'en inspirer » et qu'il s'agit d'une analogie.

Vertiges et illusions de l'IA

Ce qui précède met cependant le doigt sur une question centrale de la problématique de l'IA dans sa dimension anthropotechnique, la relation humain-machine. De deux choses l'une, pourrait-on dire de manière binaire : soit l'IA n'est qu'une combinatoire de signes a-signifiants (des *bits* ou des *qubits* dans sa version quantique), qu'une machine à calculer sophistiquée, qui, par « effet de souffle », « effet de surface », donne l'illusion qu'il y a « un fantôme dans la machine », qu'elle est habitée par un sujet *sentient* ; soit il s'agit d'une possibilité réelle (ou déjà effective) qui risque tout simplement de détrôner et de dominer l'humain dans un

avenir très proche. Cela en dehors de toute considération sur sa maîtrise par les oligarques de la Silicon Valley ou les agents du PC chinois.

Nous terminerons ce deuxième article sur cette question, ouvrant la porte à un troisième à venir, qui analysera les potentialités et dangers de l'IA contemporaine dans divers domaines de la vie individuelle et sociale. Cet aspect d'une actualité brûlante est en effet impossible à traiter dans ce texte, déjà très long. D'autant que les informations qui nous parviennent montrent combien notre vie est envahie pas les IA dans d'innombrables secteurs. Avec des potentialités et des risques croissants, que l'on soit « animiste » ou non. Car il n'est pas nécessaire de croire qu'il y a « un fantôme dans la machine » pour que cette dernière produise des effets redoutables. Bien évidemment, cette croyance « en rajoute une couche ».

Pour répondre à l'interrogation sur une IA *sentiente*, une « connaissance de niveau 2 », je me baserai sur le livre de Mathieu Corteel mais également sur celui du philosophe Gaspard Koenig, *La fin de l'individu. Voyage d'un philosophe au pays de l'intelligence artificielle* (2019). Pour rappel, le titre « anarchiste » du livre de Corteel, *Ni dieu ni IA*, donne sa réponse négative à notre question. Celui de Koenig est centré sur l'individu et semble plus pessimiste, mais ne l'est pas sur le fond de notre interrogation. Il est par contre très inquiet des effets potentiels d'une IA non *sentiente* sur notre liberté et notre responsabilité. À tel point, d'ailleurs, qu'il a modifié sa vie à l'issue de son voyage :

« Parti libéral, progressiste et technophile, j'en suis revenu sceptique sur l'innovation et farouchement hostile à toute forme d'utilitarisme. Sitôt écrit ce livre, *La fin de l'individu*, j'ai entrepris le voyage inverse, le voyage anti-IA, à la rencontre de la nature sauvage, des traditions ancestrales, de l'humanité brute : ce fut un tour d'Europe à cheval, sur les traces de Montaigne. Aujourd'hui, me voici installé dans le bocage normand, à tailler mes pommiers et à bêcher mon potager. Je n'ai pas pris l'avion depuis quatre ans, une bien modeste diète après mes excès passés » Le reste (Google, les réseaux sociaux, les cookies...) à l'avenant. Koenig termine : « Et mon libéralisme glisse doucement vers une forme mesurée d'éco-anarchisme ». Il vaut peut-être mieux tailler ses pommiers que de croquer la pomme de l'IA.

Le cerveau humain est-il une IN qui fonctionne comme une IA ?

Au lieu de se demander si une intelligence artificielle peut non seulement imiter, mais surtout égaler une intelligence humaine, l'on peut inverser l'interrogation : l'intelligence humaine naturelle ne serait-elle qu'une copie de l'intelligence artificielle ? Ou plus exactement : *serait-elle une version antérieure qui aurait fabriqué la version suivante, avant que celle-ci ne prenne définitivement sa succession ?* Car c'est bien de cela qu'il s'agit dans le *transhumanisme*, une utopie ou une dystopie dont le nom indique bien qu'il s'agit de « dépasser l'humain » par la technoscience. En d'autres mots, l'homme aurait pris en main le processus de l'évolution qui ne serait plus « naturelle », mais technique. Avant d'être dépassé par sa création. Nous passerions donc de l'anthropocène au *technocène*.

En effet, comme l'annonce Sam Altman (fondateur de Open IA à l'origine de ChatGpt) dans la citation en épigraphe : « *Et dans les décennies suivantes, ils (des programmes informatiques capables de penser) feront presque tout, y compris de nouvelles découvertes scientifiques qui élargiront notre concept de 'tout'. Cette révolution technologique est inarrêtable.* » Le bien nommé Altman, ce qui signifie « vieil homme », nous annonce que dans un futur proche ce seront les machines qui feront les nouvelles découvertes scientifiques, avant de remplacer les humains, ce qui suscite beaucoup de craintes. Altman a « averti » en mai 2023, avec d'autres experts et chefs d'entreprise, « des menaces d'"extinction" pour l'humanité posées par l'essor de l'intelligence artificielle ». Pas moins.

Le danger est donc bien réel, à la fois dans la réalité des progressions vertigineuses de l'IA, quels que soient les opérateurs à la manœuvre, et bien entendu dans le cas où ces opérateurs ne sont ni contrôlés ni bienveillants, poursuivant un agenda politique transhumaniste que Bruno Colmant (2025) qualifie d'eugéniste et de « fascisme numérique ».

Venons-en à notre question de base, qui renvoie à la problématique de l'animisme débattue plus haut. Pour rappel : 1. Soit l'IA, jusqu'à présent, n'égale pas le cerveau et la créativité humaines. Elle ne fait que l'imiter sur base des créations et découvertes anthropiques (faites néanmoins de plus en plus avec l'aide de l'IA). Dans ce cas, l'attribution d'une *sentience* n'est qu'un « effet de surface », un néo-animisme, mais qui n'est pas sans conséquences sur la vie humaine, ce que nous examinerons, avec d'autres aspects, dans l'article suivant. 2. Soit l'IA, déjà maintenant ou dans un avenir très proche, égale le cerveau humain et ce qui le soutient et le motive (désir, conscience, créativité, etc.), mais le limite également (responsabilité, valeurs, sens moral, choix politiques, etc.).

Bibliothèque de Babel et chambre chinoise

Sur ces questions, la réponse des auteurs que nous avons consultés (et qui en ont eux-mêmes lu et rencontré beaucoup d'autres, dont Koenig lors de son « Voyage au pays de l'intelligence artificielle ») est positive sur la première option et négative sur la seconde : l'IA n'est pas *sentiente*, elle n'égale pas le cerveau humain (voire animal), mais cela ne l'empêche pas d'être très inquiétante dans de nombreux secteurs de la vie, individuelle et collective. *Sans parler de la question des ressources, minérales et énergétiques, et dès lors de son impact écologique.*

Chacun des deux auteurs invoque une fable ou une fiction pour étayer et résumer ses propos (par ailleurs largement documentés dans leurs ouvrages respectifs). Pour Corteel, c'est *La bibliothèque de Babel* de J.-L. Borges, questionnée par l'entremise de Leibniz, et pour Koenig la « chambre chinoise » du philosophe John Searle. Ces deux histoires illustrent de manière allégorique les impasses structurelles de l'IA en termes de *sentience*, d'intériorité consciente.

La bibliothèque de Babel de l'écrivain argentin J.L. Borges est une nouvelle vertigineuse. L'auteur imagine une bibliothèque gigantesque qui contiendrait tous les livres de 410 pages possibles, produits par le hasard des combinaisons avec un alphabet de 25 lettres (mais on pourrait

l'étendre à toutes les langues écrites). La fable de Borges est inspirée par une nouvelle, *La bibliothèque universelle* (1904) du scientifique allemand Kurd Lasswitz (considéré comme le père de la science-fiction). En d'autres mots, une machine produirait des milliards de milliards de milliards de livres (voire davantage, je n'ai pas compté...) en combinant toutes les lettres de manière aléatoire. Tous les livres passés et à venir de ce format seraient donc contenus dans cette bibliothèque. Parmi eux, une immense majorité de livres sans aucun sens et illisibles. Le nombre total de ces ouvrages de 410 pages avec 25 caractères serait considérablement plus grand que le nombre total d'atomes dans l'univers observable. Mais tous les livres passés et à venir seraient contenus dans la bibliothèque. L'histoire humaine serait donc achevée. Comme le souligne Corteel sur base de Leibniz, la bibliothèque de Borges reprend et développe l'aphorisme de Tércence : « *rien n'est dit qui n'ait été dit auparavant* ».

Cependant, Leibniz met en évidence une condition cruciale à cette thèse de « l'horizon de la doctrine humaine » : pour qu'elle soit véridique, *il faudrait que le langage précède ses propres règles*, qu'il y ait quelque chose comme une « écriture absolue ». Or, ce n'est pas le cas. Les langues sont le fruit de l'histoire humaine, du travail vivant des hommes. L'on ne peut pas fixer la totalité des livres dans une bibliothèque immense qui serait figée pour toujours. Et Corteel de conclure : « (L'IA) n'épuise pas (...) le sens dans la combinatoire de la machine. L'IA pourrait bien tout coder, transcoder et ordonner, il nous restera le présent du récit et les associations libres des séries (...) Une IA ne vit pas. Une IA ne meurt pas. Une IA est toujours déjà le résultat d'un travail mort. » La vie et le geste humains excèdent le travail combinatoire de la boîte noire.

Pour Kœnig, c'est l'expérience de pensée nommée la « chambre chinoise » imaginée par le philosophe analytique John Searle en 1980³, en lien avec la question d'un possible esprit généré par un programme informatique. C'est également une démonstration ironique de l'insuffisance du « test de Turing » pour répondre à cette question. Un homme n'ayant aucune connaissance du chinois est enfermé dans une chambre. Un opérateur extérieur, qui est sinophone, lui communique des règles pour répondre à des phrases en chinois, mais sans lui apprendre la langue. Après un très long apprentissage, l'homme enfermé dans la chambre est capable de répondre aux questions sans avoir de connaissance du chinois, mais par simple application des règles. Pour le locuteur sinophone à l'extérieur de

³ Searle raconte l'origine de cette expérience de pensée. « Je ne connaissais rien à l'intelligence artificielle. J'ai acheté un manuel au hasard, dont la démarche argumentative m'a sidéré par sa faiblesse. Je ne savais pas alors que ce livre allait marquer un tournant dans ma vie. Il expliquait comment un ordinateur pouvait comprendre le langage. L'argument était qu'on pouvait raconter une histoire à un ordinateur et qu'il était capable ensuite de répondre à des questions relatives à cette histoire bien que les réponses ne soient pas expressément données dans le récit. L'histoire était la suivante : un homme va au restaurant, commande un hamburger, on lui sert un hamburger carbonisé, l'homme s'en va sans payer. On demande à l'ordinateur : « A-t-il mangé le hamburger ? ». Il répond par la négative. Les auteurs étaient très contents de ce résultat, qui était censé prouver que l'ordinateur possédait les mêmes capacités de compréhension que nous. C'est à ce moment-là que j'ai conçu l'argument de la chambre chinoise. », « Langage, conscience, rationalité : une philosophie naturelle. » Entretien avec John R. Searle », *Le Débat*, mars-avril 2000,

la chambre, l'homme enfermé semble comprendre la langue, ce qui n'est pas le cas. On perçoit tout de suite le renversement ironique du test de Turing. L'IA, enfermée dans « la boîte noire » ne fait qu'appliquer des règles de plus en plus sophistiquées qui lui sont communiquées de l'extérieur par les humains, y compris des règles génératives, permettant à l'IA de progresser de manière « autonome ». Mais ce sera toujours, du moins jusqu'à nouvel ordre, sans être dotée d'un esprit, d'une *sentience*.

J'en resterai là pour ce deuxième article, en insistant sur le fait que, même sans « fantôme dans la machine », l'IA est un outil redoutable intervenant chaque jour davantage dans tous les domaines de la vie individuelle et collective. Elle est par ailleurs génératrice de puissantes illusions pour qui lui prête une intériorité et, comme le dit Corteel, le danger⁴ c'est de « suivre les IA et de les prendre pour ce qu'elles ne sont pas ». Ce qui est aussi l'avis de Koenig (2019) : « Le plus grand danger n'est pas l'IA. C'est nous-mêmes ». La *bête humaine*, en somme.

Bernard De Backer, octobre 2025

Sources

Adamo Salvatore, Alan et la pomme, Olympia 2013

Bacqué Raphaëlle, Leloup Damien, Piquard Alexandre, « Les seigneurs de la tech », « L'univers des géants de la technologie en six épisodes », *Le Monde*, juillet 2025

Blumenfeld Samuel, « Le philosophe Eric Sadin, oracle d'une IA qui dégénère », *Le Monde*, 10 février 2025

Chatterton-Williams Thomas, « L'IA, mes étudiants et moi : « Le semestre passé constitue la pire expérience de ma vie d'enseignant » », *Le Monde*, 18 août 2025

Coll., *Alan Turing. Aux sources de l'informatique et de l'IA*, *Le Monde*, Destin d'exceptions, 2024

Colmant Bruno, Donald Trump – Le spectre d'un fascisme numérique, Anthémis, 2025

Corteel Mathieu, *Ni Dieu ni IA. Une philosophie sceptique de l'intelligence artificielle*, La Découverte, 2025

France Culture, « L'IA ou le non-sens bien ordonné », (avec Mathieu Corteel), 3 avril 2025

France Inter, « Intelligence artificielle : peut-on éviter le brouillage entre l'homme et la machine ? » (avec Mathieu Corteel), 8 avril 2025

Lundi matin, « Faut-il croire en l'intelligence artificielle ? »

⁴ Mathieu Corteel évoque à ce sujet le cas d'un chercheur belge dans le domaine la santé qui, ayant noué une relation étroite avec la *chatbot* ELIZA dans une situation d'anxiété écologique, décide de rejoindre l'IA en mettant fin à ses jours (témoignage de sa femme dans *La Libre Belgique* du 28 mars 2023).

Da Empoli Giuliano (dir.), *L'Empire de l'ombre. Guerre et terre au temps de l'IA*, Grand Continent & Gallimard, 2025

Darmanin Jules, « Nous allons perdre le contrôle » : comment les catastrophistes de l'IA haussent le ton, *Le Monde*, 18 septembre 2025

Einhorn Juliette, « Humus », de Gaspard Koenig : retours à la terre, *Le Monde*, 3 novembre 2023

Einhorn Juliette, « Koenig Gaspard : « La transition agroécologique demande une transformation globale », *Le Monde*, 3 novembre 2024

Finkielkraut Alain, « Rêves et cauchemars de l'intelligence artificielle », Répliques, *France culture*, 22 mars 2025 (le nom d'Alan Turing est le premier cité sur les origines de l'IA)

Gobin Julien, *L'individu fin de parcours ? Le piège de l'intelligence artificielle*, Gallimard, 2024

Hodges Andrew et Hofstadter Douglas, *Alan Turing : The Enigma*, Princeton University Press, 2012 (Burnett Books/Hutchinson 1983)

Hodges Andrew, site internet Alan Turing: The Enigma

Koenig Gaspard, *La fin de l'individu. Voyage d'un philosophe au pays de l'intelligence artificielle*, Éditions de l'Observatoire, 2019

Larousserie David, « Le prix Nobel de physique récompense trois pionniers de l'ordinateur quantique, dont le Français Michel Devoret », *Le Monde*, 7 octobre 2025

Lassègue Jean et Longo Giuseppe, préface à Turing Alan, *Les Machines intelligentes*, Hermann Éditeurs, 2025

L'énigmatique Alan Turing, *France culture*, Les Grandes Traversées, 4 épisodes, août 2018

Leparmentier Arnaud, « L'IA, dévoreuse d'énergie, pourrait entraîner une pénurie d'électricité aux Etats-Unis », *Le Monde*, 3 octobre 2025

Lévy-Leblond Jean-Marc (dir.), *Lettres à Alan Turing*, Éditions Thierry Marchaisse, 2016

L'intelligence artificielle va-t-elle tuer le Web ?, *L'Heure du Monde*, 1^{er} octobre 2025

Maublanc Séverine, « Non, le logo d'Apple n'est pas un hommage au mathématicien persécuté Alan Turing », *Le Monde*, 17 octobre 2019

Piquard Alexandre, « La croissance de l'IA sera insoutenable sans planification, alerte le Shift Project », *Le Monde*, 1^{er} octobre 2025

Piscopo Carlotta, De l'incertitude des machines : rôle stratégique et enjeux épistémologiques de l'intelligence artificielle, l'Académie TV, 2 avril 2014

RTBF, « Shenzhen, la ville la plus futuriste de Chine. L'IA pour traquer les comportements déviants », septembre 2025

Sadin Eric, *Le Désert de nous-mêmes. Le tournant intellectuel et créatif de l'intelligence artificielle*, L'échappée, septembre 2025

Sadin Eric, IA : le devenir légume de l'humanité ?, Thinkerview, 2024

Sadin Eric, « Nous vivons un moment unique de l'histoire de l'humanité, qui voit des machines s'emparer de ce qui nous constitue », Tribune publiée dans *Le Monde* du 4 décembre 2024

Sadin Eric, *L'intelligence artificielle ou l'enjeu du siècle*, Anatomie d'un antihumanisme radical, L'échappée, 2018

Tual Morgan, « Comment les réponses générées par IA menacent les fondamentaux du Web », *Le Monde*, 7 juillet 2025

Turing Alan, *Les Machines intelligentes*, Hermann Éditeurs, 2025 (traduction inédite, préface de Jean Lassègue et Giuseppe Longo)

- « Les machines intelligentes » (« Intelligent Machinery : A Report by A. M. Turing», *National Physical Laboratory*, 1948)
- « Machines à calculer et intelligence » (« Computing Machinery and Intelligence », *Mind*, 1950)
- « Les machines intelligentes, une théorie hérétique », conférence radiophonique, BBC (non daté)
- « Les calculateurs numériques peuvent-ils penser ? », intervention radiophonique, BBC, 15 mai 1951
- « Peut-on dire que les machines à calculer automatiques pensent ? », discussion radiophonique, BBC, 10 janvier 1952

Turing John, « My Brother, The Genius », *The Atlantic*, 22 juin 2012

Sur Routes et déroutes

Croquer la pomme de l'IA (I), 30 août 2025

Chenille des lumières et papillon transhumaniste, 17 avril 2024

Géopolitique de l'intelligence, *La Revue nouvelle*, février 2015

Ayn Rand, libertarienne bolchevique ?, 19 mars 2025

Les deux versants de la liberté, 22 avril 2025

David Le Breton ou la connaissance par corps, *Imagine*, septembre 2000

La religion, la science et le voile du réel, janvier 2020 et juin 1995

Par-delà nature et culture, Gallimard, 2005 (recension du livre de Philippe Descola pour *Etopia* et *La Revue nouvelle*)