



PHILOSOPHIE

« Dire que la machine va surpasser l'humanité n'a aucun sens »

En réalité, la machine est plus forte que l'homme depuis la création de la machine, confie le philosophe au « Soir ». Pour autant, nous sommes à des années-lumière de fabriquer, de toutes pièces, un artefact intelligent, un « esprit artificiel ». D'ici là, l'homme a juste le choix : se soumettre, ou pas, aux algorithmes.

ENTRETIEN
PHILIPPE LALOUX

Le bonheur est-il affaire de raison ? Vous avez deux heures. ChatGPT aura mis moins d'une minute pour se soumettre à cette épreuve de bac philo. « Mais sa copie ne valait pas un clou », relève Raphaël Enthoven qui a docilement accepté de se frotter à la machine. Cette expérience, qu'il qualifie « d'absurde », a nourri l'urgence de remettre chacun à sa place : le robot au rayon automates, l'homme dans son pré carré à jamais inaccessible par la machine, l'esprit. Dans son dernier essai, *L'esprit artificiel* (Editions de l'Observatoire, 2024), Raphaël Enthoven nous livre une réflexion sur l'intelligence artificielle au prisme de ce qui lui échappe : l'amour et la philosophie.

Pouvez-vous me prouver que vous n'êtes pas un robot ?

Je n'ai aucun effort à faire pour ça puisque la façon que j'ai de réagir à cette question n'est pas une déduction. Elle relève de ce qu'on appelle la « spontanéité » : c'est une réaction qui n'épuise pas ce qu'il y a d'original dans ma façon de réagir. C'est en cela que mes réponses ne peuvent pas être synthétisées par une machine. Elles peuvent être imitées ou reproduites après coup. Mais en aucun cas une machine ne peut réagir comme je le fais.

Paradoxalement, vous avez intitulé votre essai *L'esprit artificiel*, mais, sans en « divulguer » l'issue, vous consacrez 180 pages à démontrer qu'en aucun cas l'esprit peut être artificiel...

Le titre est une fausse piste, une façon de montrer l'absurdité de la chose. L'idée du titre m'est venue en partie du livre d'Henri Atlan, *L'utérus artificiel*. Sauf que dans cet ouvrage – qui était une fiction scientifique – Atlan présentait l'utérus artificiel comme devant advenir et techniquement réalisable. Alors que pour moi, pas du tout : l'esprit artificiel, ce n'est même pas simplement un oxymore, soit deux termes qui s'opposent apparemment mais qui vivent ensemble, mais une contradiction profonde. Ça ne veut rien dire, en fait. Voilà. C'est un titre qui *veut ne rien dire*.

Si l'esprit ne peut pas être artificiel, qu'est-il donc ?

Rien n'est moins définissable, rien n'est plus important. D'abord, on peut dire ce qu'il n'est pas. Ni la mécanique, ni la science exacte, ni la science hypothéticodéductive, mais l'exigence qui donne le jour à tout cela. L'esprit désigne l'inquantifiable au sein de la vie. Ce dont on dit communément que « ça n'a pas de prix ». Ou bien ce qu'on n'explique pas. Par exemple, si vous regardez un coucher de soleil, vous allez trouver ça beau. Et vous allez présumer que votre émotion est universelle puisque vous direz « c'est beau » (et non « je trouve ça beau »). Vous parlerez au nom du monde entier. Pourtant, votre propos n'est pas d'imposer votre opinion. Et vous ne savez pas ce que c'est que la beauté, mais vous savez ce qui est beau. Vous reconnaissez la beauté sans la connaître. Autrement

dit, nous dit Kant, l'expérience du beau est une expérience à la fois universelle et sans concept. J'appelle « esprit » la capacité que nous avons à faire des expériences universelles qu'aucun concept ne saisit ou ne résume. J'appelle esprit la singularité, l'irréductibilité de l'étoffe humaine à toute entreprise de synthèse.

Vous écrivez que, en philosophie, l'IA ne sert finalement à rien...

L'IA ne sert à rien parce qu'il y a ce qui nous fait penser, mais d'abord ce qui nous impose de penser, c'est-à-dire l'expérience d'un grand désarroi. Le désarroi d'être là, la stupeur d'être né par hasard... et la certitude d'avoir à mourir. Soit quantité de choses qu'aucune machine n'éprouve. Ce qui nous dicte l'urgence de penser, ce n'est pas seulement le besoin de connaître. C'est tout ce qui, avec le besoin de connaître, permet la familiarisation d'un monde qui, par cette grâce, devient plus vivable par et pour nous. A l'origine de l'urgence de penser, on trouve des éléments qu'aucune machine ne synthétise, qu'aucune machine ne connaît, qu'aucune machine ne ressent. C'est la raison pour laquelle on ne peut pas dicter à une machine de réfléchir. Pas plus qu'on ne peut dicter au génie d'Aladin de ressusciter les morts ou de provoquer l'amour. Ces deux limites essentielles sont exactement celles de l'IA.

Le fait que la machine apprenne, ce qui la rend toujours plus efficace, ne la rend pas humaine pour autant

”

De ce fait, dites-vous, l'humanité est vouée à demeurer un casse-tête pour la machine...

Oui. Dire que la machine va s'humaniser, reproduire l'humanité ou la dépasser, n'a strictement aucun sens. D'une part, parce que c'est déjà fait. La machine est plus forte que l'homme depuis la création de la machine, c'est même sa raison d'être. D'autre part, parce que l'humanité demeure un casse-tête pour la machine. Croire qu'une machine peut devenir humaine, c'est chercher au mauvais endroit, c'est chercher du côté du calcul, ce qui n'est pas une affaire de complexité, mais de simplicité. L'humanité est aussi loin de la machine que le cœur est loin de la raison. Nous sommes à l'autre bout de l'univers de fabriquer un jour, de toutes pièces, un artefact intelligent.

Ce qui est intéressant, c'est l'illusion qu'on peut y parvenir. C'est ce vieux rêve qui traverse l'histoire de l'humanité, du Golem à l'IA. Le sentiment que l'humanité peut faire l'objet d'une synthèse repose sur l'illusion qu'un individu est réductible à la totalité des traces qu'il laisse. Or, celles-ci ne forment qu'un amas de feuilles mortes, un tas de cendres. Il y a entre les traces que nous laissons et ce que nous sommes le même écart qu'entre un tas de cendres et un incendie. Et l'on ne fait pas un incendie avec un tas de cendres. Ni avec de l'eau.

Est-ce pour le démontrer que vous avez accepté de vous confronter à ChatGPT à l'épreuve du bac de philo ?

Je suis un honnête homme et j'aime à vérifier ce que je pense. Je voulais m'assurer que la machine était parfaitement inapte à faire de la philosophie.

Nous ne sommes donc pas dans le même registre que le match historique entre Kasparov et Deep Blue ?

Non, cela n'a rien à voir. La comparaison est flatteuse, mais absurde. D'abord, Kasparov est un génie. Ensuite, Deep Blue est une extraordinaire usine à jouer aux échecs. Compilant la totalité des parties des grands maîtres, Deep Blue disposait d'une mémoire fatalement supérieure à celle du prodigieux Kasparov. La victoire de Deep Blue est une victoire qui ne renseigne pas sur la faiblesse de l'humanité. C'est une victoire qui renseigne sur le fait qu'une usine est plus forte qu'un homme. Sur le fait que le métal est plus solide que le muscle. Tout simplement. Contrairement à Deep Blue qui est une extraordinaire machine à jouer aux échecs, ChatGPT est absolument, durablement, définitivement, inapte à faire de la philosophie. Donc il n'y avait pas de match. Ce n'est pas comme s'il y avait eu deux philosophes qui s'affrontent, deux personnes de bonne volonté qui mettent leurs connaissances en partage. C'était une machine débile qui ne comprend rien à l'exercice en question, face à un prof de philo.

Vous soulevez les limites de l'IA à l'égard de la philo. Pourrait-on pointer les mêmes limites à l'égard de l'art ?

L'IA peut reproduire à l'infini des chefs-d'œuvre « à la façon de ». Autrement dit, l'IA est un merveilleux faussaire. Mais qu'est-ce qui fait que le faussaire, de son propre aveu, n'est pas un artiste, alors qu'il est parfaitement capable de reproduire les toiles de maître ? Parce qu'une œuvre d'art, qui obéit à une série causale qu'on serait incapable de synthétiser, d'analyser et de reproduire, est traversée par le temps. C'est ce qui fait dire à Kant qu'Homère ne peut pas expliquer d'où lui viennent les vers de l'Odyssée, alors que Newton peut expliquer d'où lui viennent ses équations. La science est traversée par l'éternité, elle est donc réversible. L'art est traversé par le temps. Il n'est pas réversible, pas plus qu'une mayonnaise. C'est la raison pour laquelle une œuvre d'art porte en elle des normes nouvelles. Quand Corneille écrit *Le Cid*, il fait exploser les règles de l'art classique, mais il crée aussi une norme nouvelle qu'on peut ensuite reproduire à l'infini. L'IA intervient dans un second temps. Une IA peut à l'infini pisser du Corneille et du sous-Corneille, mais en aucune manière avoir l'intuition d'une pièce qui déjoue les règles de l'art classique de telle manière qu'elle produise elle-même une norme nouvelle. C'est la raison pour laquelle, juste au passage, il n'y a évidemment pas d'histoire de l'art. De la même manière qu'il n'y a pas d'histoire des idées. Parce qu'on ne peut pas considérer qu'un artiste ou un penseur progresse par rapport à un autre. Malevitch n'est pas meilleur peintre que Léonard de Vinci. Kant n'est pas meilleur en philosophie que Leibniz. Leibniz n'est pas meilleur que Platon. La philosophie ne progresse pas, l'art non plus.

Quand OpenAI a lâché son jouet, ChatGPT, entre les mains du grand public, certains ont entretenu cette illusion d'une machine à tout faire, y compris penser à notre place...

La confusion est dans le sentiment qu'il suffirait de la recommandation adéquate pour que la machine produise de la réflexion. Idée confuse

Raphaël Enthoven

Né en 1975, ancien étudiant de l'École normale supérieure (ENS) et agrégé de philosophie, ambassadeur, Raphaël Enthoven a depuis toujours baigné dans un milieu intellectuel privilégié. Fils de l'homme d'affaires Jean-Paul Enthoven, il enseigne successivement à l'Université Lyon 3, à l'ENS de Lyon, ainsi qu'à l'Institut d'études politiques de Paris. Il est aussi le cofondateur, avec Michel Onfray, de l'Université populaire de Caen. Auteur prolifique, le philosophe est aussi un homme de médias (France Culture, Arte, *Philosophie Magazine*...).

PH.L.



L'esprit artificiel
RAPHAËL ENTHOVEN
Editions de l'Observatoire
185 pp.
19 euros

Une semaine au cœur de l'IA avec « Le Soir »

On dit d'elle qu'elle est la 4e révolution industrielle, trois décennies à peine après la précédente, celle d'internet. Agée pourtant d'une cinquantaine d'années, elle a refait irruption dans nos vies de manière fracassante à la lumière de quelques « robots conversationnels » aux performances troublantes, qui fascinent autant qu'elles effrayent. OK. Mais encore... Comment le monde de l'entreprise, le secteur de la santé, celui de la mobilité, le monde de l'enseignement et l'univers de la création artistique se sont emparés de l'IA ? Les questions qu'elle soulève, les espoirs qu'elle fait naître, aussi, pour l'avenir de l'humanité. De ce samedi à vendredi, la rédaction du *Soir* propose de vous plonger au cœur des enjeux et des impacts concrets de ces technologies. PH.L.

entre toutes. Comme si la machine était un génie docile, spontanément doté des qualités qu'on lui ordonne d'avoir... Ils ont travaillé le « prompt », soit l'ensemble des recommandations adressées à la machine, pendant plusieurs jours avec un prof de philo. Mais comme ces (excellentes) recommandations exogènes qui lui étaient soumises ne procédaient pas d'un exercice de réflexion, elles n'avaient aucune chance d'aboutir. Aucune et durablement. Et le perfectionnement de ChatGPT 4, 5, 7, 12... ne le rend pas davantage perméable à cette façon de faire. Il y a là, une fois de plus, quelque chose d'irréductible.

Y a-t-il, dans la copie remise par ChatGPT, des éléments qui vous ont malgré tout impressionnés ? Pas mal de lycéens rêveraient de décrocher un 11/20 à l'épreuve du bac de philo...

Non. Sa copie est nulle. C'est irrécupérable. Ce qui est intéressant, c'est que ChatGPT a réagi comme un mauvais élève. Il a vu la notion de « bonheur », celle de « raison », et s'est demandé ce qu'il savait sur ces affaires-là. Il est allé chercher toutes les théories qui leur correspondaient. Les théories en question, il les avait toutes à sa disposition. Il les a ensuite synthétisées dans trois tiroirs indifférents. Le troisième aurait pu venir avant le premier. Il n'y avait aucune nécessité à ce qu'il disait, aucune problématique, aucun exercice de réflexion. Si impressionnant soit-elle par la quantité de données dont elle dispose, cette copie ne mérite pas la moyenne parce que ce n'est pas là-dessus qu'on évalue un élève. En philosophie, on évalue sa capacité à se poser des questions, à réfléchir. Or, précisément, la machine en est absolument incapable.

Et en même temps, reconnaissons que si on nous demandait, à vous ou à moi, de séquencer le génome humain en quelques minutes, on n'y arriverait pas...

Encore une fois, je ne dis pas que les outils ne sont pas plus forts que nous. Bien sûr, ils le sont. C'est même leur raison d'être. Heureusement qu'une machine est capable de faire ce que nous ne pouvons pas faire. Sinon, à quoi bon ? Mais reste qu'elle n'en a pas l'initiative. Une fois qu'on a séquencé le génome de quelqu'un, ce que l'on sait de lui (la possibilité d'un cancer, la potentialité de tel ou tel problème, etc.) n'est pas l'essentiel de la per-