

La quête d'intelligences artificielles conscientes soulève de multiples questions éthiques. Etat des lieux

Les robots auront-ils bientôt une âme?

« ANNE-SYLVIE SPRENGER
PROTESTINFO

Etats d'âme (1/3) » Entre les questions sur la moralité des algorithmes de ChatGPT et les promesses des deadbots – ces robots conversationnels permettant de faire parler les morts –, la quête d'intelligences artificielles conscientes ne cesse de soulever des questions éthiques. Mais où en est réellement l'état des avancées? A quel degré d'humanité les robots de demain pourront-ils prétendre? A quelles conscience, émotions, spiritualité auront-ils un jour accès? Mise au point avec Ezekiel Kwetchi Takam, doctorant en éthique théologique à l'Université de Genève, spécialisé dans les enjeux de l'intelligence artificielle (AI).

De nombreuses entreprises, Google en tête, promettent le développement de robots pourvus d'une conscience. Où en sont réellement les avancées en la matière?

Ezekiel Kwetchi Takam: C'est vrai que dans la continuité des thèses de la singularité technologique, portées en l'occurrence par Ray Kurzweil (ex-directeur de l'ingénierie chez Google), la quête des intelligences artificielles conscientes reste toujours un sujet d'actualité. Il faut avouer cependant que les avancées vers cet «exploit» sont encore au stade embryonnaire. Blake Lemoine, ex-ingénieur chez Google, avait défrayé la chronique en annonçant médiatiquement que le Chatbot LaMDA, dont il avait la charge d'entraînement, avait acquis une conscience sentiente (capable de sensations, ndlr). Suite à cette information, le simple fait que Google se soit désolidarisé de cet ingénieur en dit long sur les actuels enjeux stratégiques des producteurs d'IA.

C'est-à-dire?

Aujourd'hui, les courses hégémoniques vers la domination du marché de l'IA invitent les acteurs à être beaucoup plus pragmatiques, moins expérimentaux, et plus focalisés dans des objectifs d'efficacité, et ceci dans des domaines précis. Cela dit, au regard de cette question de l'IA consciente qui attise une bonne partie du débat sur le développement de l'IA, il me semble urgent et important de démystifier cette caractéristique par un travail de définition étymologique.

Justement, qu'entend-on par les termes «âme» ou «conscience»?

Pour la science, l'âme est considérée comme étant une donnée métaphysique non prouvable empiriquement. En théologie, elle garde toujours une valeur capitale dans la définition du vivant et de sa relation à son créateur. Quant à la conscience, elle renvoie étymologiquement à l'«échange de connaissances morales avec un autre», cet autre qui

nous aide à forger notre conception du bien et du mal. Et tandis que la théologie chrétienne identifiera l'«autre» à Dieu, de son côté, la science, qui découle de la philosophie, identifiera l'«autre» à la raison, qui peut elle-même dépendre d'une culture. Dans cette considération, tout comme un nouveau-né acquiert la conscience durant son développement grâce à ses échanges avec

son environnement, l'IA peut aussi développer et partager des connaissances morales, et donc acquérir une certaine conscience dans son rapport et son échange avec son environnement.

Une intelligence artificielle peut donc être capable de jugements moraux?

Absolument. C'est ainsi que ChatGPT a pu évoluer d'une application refermant certains biais sexistes et racistes en novembre 2022 vers une application qualifiée (ou critiquée en fonction des angles d'appréciation) de woke dès fin février 2023. Simplement parce qu'au cours de ses quelques mois d'existence, il a pu apprendre de ses erreurs, de ses échanges, et développer, dans un esprit mimétique, un prisme moral qui se rapproche du nôtre. Au même titre qu'un nouveau-né développerait une conscience en mimant l'entourage qui encadre son développement.

Et qu'en est-il de l'âme?

N'est-ce pas pourtant ce que certaines entreprises cherchent à capter avec ces robots conversationnels qui proposent de ressusciter virtuellement nos défunts?

En effet, on constate qu'à l'ère du numérique, une forme d'identité individuelle (informationnelle et narrative) peut être traduite en data, du fait de notre présence dans l'univers du digital et de l'algorithmique. C'est cette double identité informationnelle et narrative qui est actualisée et vivifiée par ces nouveaux robots-IA conversationnels.

Des robots pourvus d'une âme, cela vous semble-t-il réaliste?

A cette question, Marvin Minsky, l'un des pères fondateurs de l'IA, affirmait en mai 2014, dans une interview accordée au journal *The Jerusalem Post*, que les IA pourront avoir, dans les années à venir, une âme. Cependant, sa réponse décentrait le sens de l'âme, de sa conception religieuse (souffle de vie) vers les notions

de l'identité et de la destinée (ou simplement du but). Il entend par âme cette idée du «qui nous sommes» et du «pour quoi nous sommes». Et dans cette considération, sachant que la particularité des IA est leur ouverture aux outputs – c'est-à-dire leur capacité d'échange d'informations avec l'environnement extérieur –, il ne serait pas impossible dans les années à venir que les IA interconnectées puissent communiquer, constituer un écosystème identitaire, et se définir un projet commun. Cependant, il me semble difficile de parler ici de l'«âme» qui, à mon sens, est encore le monopole du vivant.



«L'intelligence artificielle peut développer des connaissances morales» Ezekiel Kwetchi Takam

Justement, qu'en penser d'un point de vue théologique?

Théologiquement, c'est d'autant plus intenable. Si l'on revient aux premières occurrences de l'âme dans le récit biblique, l'expression employée est *Nephesh*, qui se traduirait littéralement par être vivant, sans aucune distinction entre les humains et les non-humains. Et ce vivant, de ma modeste observation, est caractérisé par deux choses. Premièrement, le désir, qui instaure en lui un mouvement vers une finalité. Et deuxièmement la vulnérabilité, c'est-à-dire sa fragilité et la possibilité de voir sa volonté être obstruée par des forces extérieures ou intérieures à lui. C'est ce principe de vulnérabilité qui donne d'ailleurs toute sa pertinence et tout son sens à la notion du salut: Dieu, se faisant chair, accompagne le vivant dans son expérience de vulnérabilité, en vue de l'en extraire et de lui proposer un horizon de paix, de joie, etc. Une fois que nous pourrions démontrer que les IA peuvent avoir ce désir et cette vulnérabilité, elles pourront être considérées comme détentrices d'âme. Pour l'instant, ça me semble être une impossibilité. »

» Notre série d'été interroge trois états d'âme propres au XXI^e siècle, entre les promesses de l'intelligence artificielle, les affirmations antisécularistes accordant une vie intérieure aux animaux et l'importance toujours plus reconnue de la spiritualité dans les soins.



Adobe Stock

Caméra cachée en Terre sainte

Actes antichrétiens » Pour enquêter sur la multiplication des gestes antichrétiens dans la ville, pour la chaîne israélienne Yossi Levi s'est immergé dans le quotidien d'un franciscain à Jérusalem. Il ne lui a pas fallu plus de cinq minutes dans la Vieille-Ville pour recevoir insultes et crachats.

Habillé de la bure franciscaine, le journaliste Yossi Levi était accompagné du Frère Alberto Pari, secrétaire de la cus-

todie de Terre sainte, et enquêtait sur la recrudescence des actes antichrétiens dans la ville, pour la chaîne israélienne Channel 13, rapporte le site terresainte.net. Pour que les téléspectateurs israéliens prennent la mesure des incivilités quotidiennes dont les prêtres chrétiens peuvent être la cible, la custodie avait joué le jeu. Elle avait prêté un habit au journaliste qui y a dissimulé une petite caméra. » CATH.CH

SENIORS

INDULGENCE PLÉNIÈRE

Le 23 juillet, l'Église catholique accordera une indulgence plénière à tous les fidèles qui participeront à la 2^e Journée mondiale des grands-parents et des personnes âgées, et qui seront «animés d'un véritable esprit de pénitence et de charité». Cette journée donnera lieu à une messe présidée par le pape François à la basilique Saint-Pierre. CATH.CH

Martyrs du siècle à répertorier

Cause des saints » En vue du Jubilé 2025, le pape institue une commission des nouveaux martyrs du XXI^e siècle.

L'objectif de cette structure d'une quinzaine de membres sera d'élaborer un «catalogue» de chrétiens de toutes confessions qui sont morts pour leur foi, de l'an 2000 à nos jours, sans pour autant que leur cause de béatification soit ouverte à

Rome. Cette commission, précise le pape dans sa lettre, «n'entend pas établir de nouveaux critères pour la vérification canonique du martyre». Elle devra répertorier les noms de «ceux qui, aujourd'hui, continuent à être tués seulement parce que chrétiens».

Les travaux historiques permettront de dresser une liste de baptisés qui ont vécu le martyre chrétien, mais ne sont pas re-

connus officiellement martyrs par l'Église. Sans ce recensement, ces chrétiens resteraient anonymes.

Le cardinal Kurt Koch, préfet du Dicastère pour la promotion de l'unité des chrétiens, a salué l'œcuménisme de cette commission. «Je pense qu'il s'agit d'une grande confirmation de l'œcuménisme des martyrs», a-t-il indiqué à *L'Osservatore Romano*. » CATH.CH