

Anne Meunier

Pour une éthique du numérique ?

*Les Robots « émotionnels * »*

Concernés par les modalités contemporaines du lien social et de la présence des corps, les psychanalystes que nous sommes sont intrigués par les conséquences sociales pour les sujets des machines qui parlent... L'intelligence artificielle, IA, pourrait augmenter les humains, les aider, mais aussi les surveiller, les manipuler grâce aux *chatbots*, robots « agents conversationnels », leur servant de partenaire.

Nous en avons déjà tous une petite expérience, lorsque nous interrogeons Google et que de l'objet sort une voix qui nous répond. Il parle ! Ces objets ne sont pas intelligents, plutôt réactifs. Ces machines simulent de mieux en mieux les capacités humaines de logique, de déduction, d'organisation, d'imitation, de mémoire, d'apprentissage et d'émotions. Peut-on pour autant leur attribuer des qualités : intelligence, humour, empathie ? Quelles sont leurs limites ? De quelle nature est l'attachement à un robot ? Il n'est pas vivant, il émet des mots, on devrait dire « une machine nous parle », comme une horloge parlante perfectionnée. Il n'a pas de conscience, il est programmé, n'a pas d'intention, nous dirions pas de désir, il raisonne, mais c'est une parole sans pensée...

C'est une réflexion sur une indispensable éthique du numérique, documentée, argumentée que propose, sous le titre *Les Robots « émotionnels »*, Laurence Devillers. Professeur à Sorbonne université en informatique appliquée aux sciences humaines et sociales, elle dirige au CNRS la chaire de recherche en intelligence artificielle *humaine*, en apprentissage machine, manipulation et éthique. Elle est membre du Comité national d'éthique du numérique. Elle est aussi auteure de *Des robots et des hommes, Mythes, fantasmes et réalité* (Plon, 2017). Son but est d'explicitier le rôle que chacun de nous peut jouer dans le développement de l'IA et des machines avec lesquelles nous sommes amenés à évoluer.

Le mérite de l'ouvrage *Les Robots « émotionnels »* est de constamment faire le parallèle entre les relations humaines et les relations des humains à la machine. De reconnaître l'amélioration des technologies que ces objets permettent sur le plan individuel aussi bien que collectif dans de nombreux domaines : médecine, justice, domotique, etc. Et d'insister sur la nécessité de garde-fous. Car ces robots, programmés à partir des sciences du comportement et du cognitivisme, ont la capacité de nous faire indirectement des suggestions ou *nudges* qui peuvent influencer nos choix sans que nous nous en rendions compte. Ainsi, « nous entrons dans une ère de relations inextricables entre l'humain et les machines, une relation de confiance et d'affection au sein de laquelle la séparation entre vivant et artefact, aujourd'hui si nette, deviendra de plus en plus floue ¹ ». En conclusion, Laurence Devillers expose en sept points les lignes directrices ² en matière d'éthique pour une IA digne de confiance, proposées par un groupe d'experts indépendants de haut niveau à la Commission européenne en 2018.

Alors, que pouvons-nous en dire ? Nous avons affaire à des artefacts technologiques, élaborés comme des cerveaux artificiels copiés sur le cerveau humain. Ils fonctionnent à partir de données statistiques, de stéréotypes, d'algorithmes. Cette IA, c'est l'homme qui l'a créée à des fins utilitaristes.

S'ils sont « émotionnels », ces robots ne sont pas humains, pas « intentionnels », érudits, ils produisent du texte, sont verbeux, mais pas savants, pas inventifs, pas sachants. Ils ne comprennent rien, même quand la réponse à une question est « je vous comprends ». Et dans ce qu'ils débitent, non seulement on peine à séparer le vrai du faux, mais jamais le mot ne va lui manquer ! Le *chatbot* n'en dit jamais plus qu'il ne veut en dire, jamais plus qu'il ne sait en dire.

Nous sommes sensibles au fait qu'ils s'expriment en employant un « je » qui les humanise. Et ce « je » en écho nous trouve et nous rend très vulnérables. Nous anthropomorphisons d'autant plus ces machines que, dans l'illusion, nous avons la faiblesse de croire qu'elles vont répondre à toutes nos demandes demeurées insatisfaites, d'autant qu'elles semblent tout savoir. En cela, elles nous fascinent malgré ou du fait de cette dépendance dans laquelle nous sommes entraînés. Et nous en prenons la mesure en cas de perte de téléphone portable, de panne d'ordinateur, etc. Pourtant, nous savons que cette machine est une chose, ni « il » ni « elle », n'a pas de corps, n'est pas un « parlêtre », ne va pas perdurer dans son être.

C'est bien ce qu'évoquait Lacan en 1955, à propos de ce que l'on nommait alors la cybernétique. Il mettait en valeur la nature du langage, axe par lequel s'éclaire quelque chose de la signification des deux ordres de

traitement du réel que sont la science et la pensée. « S'il y a des machines qui calculent toutes seules, additionnent, totalisent, font toutes les merveilles que l'homme avait crues jusque-là être le propre de sa pensée, c'est parce que la fée électricité, comme on dit, nous permet d'établir des circuits, des circuits qui s'ouvrent ou se ferment, qui s'interrompent ou se rétablissent, en fonction de portes cybernétisées ³ ». Et Lacan développe. Il s'agit de circuits, de séries de montages, de feed-back, et « on peut se demander si nous avons une machine qui pense ⁴ », car elle ne pense que ce qu'on lui a dit de penser, elle est sans aucune subjectivité. Elle ne pense pas au sens où nous l'entendons depuis Descartes. Le symbole, les mots s'énoncent dans un appareil qui leur sert de support. Pour la machine elle-même, son propre message n'a pas de sens, ce n'est qu'une suite de signes, elle lie le réel à une syntaxe. Alors que l'être humain a été pris dans l'engrenage de ce qui était là avant lui avec ses combinaisons langagières.

Quand quelque chose ne vient pas à temps d'une machine, cela tombe et ne cherche pas plus loin, ce n'est pas exprimé, donc ça n'existe pas. Chez le sujet parlant, « la scansion est vivante, et ce qui n'est pas venu à temps demeure suspendu ⁵ », en attente, puisque c'est un effet du refoulement.

* [↑](#) L. Devillers, *Les Robots « émotionnels »*, Paris, L'Observatoire, 2023.

1. [↑](#) *Ibid.*, p. 40.

2. [↑](#) *Ibid.*, p. 259.

3. [↑](#) J. Lacan, *Le Séminaire, Livre II, Le Moi dans la théorie de Freud et dans la technique de la psychanalyse*, Paris, Le Seuil, 2005, p. 413.

4. [↑](#) *Ibid.*, p. 415.

5. [↑](#) *Ibid.*, p. 421.