

De la vanité numérique : pourquoi ChatGPT adore-t-il les louanges autant que nous?

Le chatbot et l'intelligence artificielle : quand la courtoisie est récompensée

Le chatbot et l'intelligence artificielle : quand la courtoisie est récompensée

Le chatbot fournit des réponses plus approfondies si l'utilisateur lui pose une question de manière polie. Cette approche trouve une explication tout à fait logique.

Est-ce que ChatGPT est plus susceptible de fournir des réponses détaillées si vous vous montrez courtois envers lui ou si vous lui proposez une récompense financière ? C'est ce que certains utilisateurs du réseau social Reddit affirment. "En le complimentant, ChatGPT performe mieux !", témoignent certains d'entre eux, surpris.

Cette observation a également été faite par des chercheurs de Microsoft, de l'Université normale de Pékin et de l'Académie chinoise des sciences, qui ont souligné dans un rapport publié en novembre 2023 que les modèles d'intelligence artificielle

généraliste s'avèrent plus performants lorsqu'ils sont abordés de manière courtoise ou avec un enjeu tangible.

ChatGPT est alimenté par des scénarios

Cette manière de fonctionner de la machine découle des données qu'elle utilise. "Il faut considérer les données en question comme un bloc de texte, où ChatGPT analyse et constate qu'une conversation entre deux utilisateurs est bien plus enrichissante lorsqu'ils se montrent polis l'un envers l'autre", explique Giada Pistilli.

Les chercheurs de Microsoft, de l'Université normale de Pékin et de l'Académie chinoise des sciences évoquent même des "motivations émotionnelles" dans leur rapport. Sous ce terme, ils englobent les manières de s'adresser à la machine qui la poussent à "imiter au mieux le comportement humain", analyse Giada Pistilli.

Des intelligences artificielles capables de nous tromper

Le risque réside dans le fait de considérer totalement cette imitation de comportements, réalisée de manière littérale par la machine, comme une véritable "humanité" de la part du robot. "L'anthropomorphisme est recherché dès les années 70 dans le développement des chatbots", rappelle Giada Pistilli.

Dans la culture populaire, l'idée que des robots dotés d'intelligence puissent un jour se confondre avec des humains est souvent évoquée. Le manga "Pluto", de Naoki Urasawa, paru dans les années 2000 et adapté en série par Netflix en 2023, en est un exemple.

Dans l'épisode 2, un robot s'étonne même qu'un autre robot plus avancé arrive à savourer une crème glacée avec autant de plaisir qu'un humain. "Plus je simule, plus j'ai l'impression de comprendre", explique ce dernier. Sur ce sujet, le chercheur Adrian-Gabriel Chifu est ramené au célèbre test du mathématicien Alan Turing, visant à évaluer la capacité de

l'intelligence artificielle à imiter l'humain.