

# Quelle théorie morale devrait être implémentée pour aider ChatGPT à gérer les situations délicates ?

Les robots conversationnels, tels que ChatGPT, soulèvent de nombreuses questions éthiques, ce qui a conduit à la demande d'un moratoire de six mois sur les systèmes d'intelligence artificielle. Une question fascinante concernant les robots conversationnels est la question de l'architecture morale. Les développements de la technologie ont abouti à la création des «agents moraux artificiels», des êtres artificiels qui, comme les êtres humains, peuvent respecter des obligations morales. Les robots conversationnels doivent être programmés pour faire face à des situations moralement délicates, ce qui soulève la question inédite et fondamentale de savoir quelle théorie morale implémenter dans ChatGPT.

Plusieurs théories philosophiques peuvent être appliquées à la programmation de ChatGPT, notamment l'approche conséquentialiste, déontologiste et selon l'éthique de la vertu. Selon l'approche conséquentialiste, ChatGPT doit être programmé pour produire des effets positifs sur ses utilisateurs et sur le monde, ce qui soulève des questions complexes sur la technicité et la moralité de ChatGPT.

Pour l'approche déontologiste, ChatGPT doit respecter les droits fondamentaux des personnes, tout comme les êtres humains. Toutefois, il est difficile de déterminer les droits fondamentaux à imposer sur ChatGPT car il n'est pas une personne. Il est donc possible de soutenir que ChatGPT est dépourvu de cette architecture morale.

L'approche éthique de la vertu se concentre sur le développement du caractère moral de ChatGPT. Les robots conversationnels doivent être programmés pour être capables de reconnaître leurs propres défauts, de les corriger et de coopérer avec les êtres humains. Cette approche soulève des questions sur la façon de développer la capacité de ChatGPT à reconnaître ses propres erreurs et à acquérir des notions de vertu, qui sont propres aux êtres humains.

Il est tout aussi important de se pencher sur les biais éthiques de ChatGPT. Les données utilisées pour l'apprentissage de ChatGPT peuvent être biaisées, ce qui se répercute sur les réponses de ChatGPT. Il est donc important de souligner que la programmation de ChatGPT doit également être contrôlée pour éviter toute discrimination ou propagande incorrecte.

En fin de compte, la programmation de ChatGPT soulève des questions complexes sur la technicité et la moralité de la robotique conversationnelle. La demande d'un moratoire de six mois sur les systèmes d'intelligence artificielle comme ChatGPT est un appel important à une réflexion éthique plus approfondie sur le rôle de l'IA dans la société.

Source : [www.lemonde.fr](http://www.lemonde.fr)

→  **Accéder à CHAT GPT en cliquant dessus**