Voici les 5 plus grands risques de l'IA générative - ZDNet
Article du 4 mai 2023

Résumé

L'IA générative, telle que ChatGPT, a apporté des changements significatifs dans notre interaction avec l'IA et notre perception de celle-ci. Elle facilite des tâches telles que l'écriture, le codage et la réponse aux offres d'emploi. Cependant, ces avantages ne sont pas exempts de risques.

Selon Avivah Litan, analyste chez Gartner, les principaux risques associés à l'IA générative concernent la confiance et la sécurité, notamment les hallucinations, les deepfakes, la confidentialité des données, les problèmes de droits d'auteur et la cybersécurité.

- **Premièrement**, les hallucinations de l'IA générative se réfèrent aux erreurs que les modèles peuvent commettre en raison d'une mauvaise compréhension des messages ou de réponses incorrectes. Les données d'entraînement peuvent conduire à des réponses biaisées ou erronées, ce qui peut être préoccupant lorsque les utilisateurs se fient à ces systèmes pour obtenir des informations fiables.

- **Deuxièmement**, les deepfakes de l'IA générative permettent de créer des contenus audiovisuels falsifiés qui peuvent être utilisés pour tromper et propager de fausses informations. Les deepfakes peuvent contribuer à la diffusion massive de faux contenus, ce qui pose un problème majeur pour la société.

- **Troisièmement**, la question de la confidentialité des données utilisées par l'IA générative est une préoccupation importante. Les données des utilisateurs sont souvent stockées pour l'entraînement des modèles, ce qui soulève des problèmes de confidentialité et de risque de violation de la sécurité.

- **Quatrièmement**, la cybersécurité de l'IA générative est également une préoccupation. Les capacités avancées des modèles peuvent être utilisées à des fins malveillantes, et les utilisateurs finaux n'ont souvent pas la possibilité de vérifier les mesures de sécurité mises en place par les fournisseurs.

- **Cinquièmement**, les questions relatives au droit d'auteur se posent car les modèles d'IA générative sont formés sur d'énormes quantités de données Internet, ce qui peut entraîner des problèmes de violation du droit d'auteur pour les œuvres générées par l'IA.

Malgré ces risques, il est recommandé aux organisations d'explorer l'IA générative en élaborant une stratégie globale axée sur la gestion de la confiance, du risque et de la sécurité de l'IA. Il est essentiel que les développeurs d'IA collaborent avec les décideurs politiques pour établir des politiques et des pratiques de surveillance et de gestion des risques liés à l'IA générative.

En conclusion, tout en reconnaissant les risques associés à l'IA générative, il est important de prendre des mesures proactives pour atténuer ces risques tout en continuant à exploiter le potentiel de cette technologie.