

Continue





























utilisant une combinaison de deux sources principales :Un volume massif de textes échantillons extraits de pages web et de code de programmation.Des conversations fournies par de vrais humains, qui ont démontré les réponses souhaitées aux prompts fournis, puis ont classé les résultats du modèle selon la qualité de la réponse Fournir des retours occasionnels d'humains à un modèle d'IA est une technique connue sous le nom d'apprentissage par renforcement à partir de feedback humain (RLHF). L'utilisation de cette technique peut aider à affiner un modèle en améliorant à la fois la sécurité et la fiabilité.Fonctionnalités et limites de ChatGPT (et autres IA génératives)ChatGPT représente une avancée passionnante dans l'IA générative, avec plusieurs fonctionnalités qui pourraient aider à accélérer certaines tâches lorsqu'il est utilisé de manière réfléchie. Il présente également des limites. Comprendre à la fois les fonctionnalités et les limites est essentiel pour tirer le meilleur parti de cette technologie.Fonctionnalités de ChatGPTParmi les fonctionnalités de ChatGPT, on trouve la capacité à :Générer du texte semblable à celui des humains qui imite le style et la structure des données d'entréeGénérer une réponse à un prompt ou un texte d'entrée donné. Cela peut inclure l'écriture d'une histoire ou la réponse à une question.Générer du texte en plusieurs languesModifier le style du texte généré (formel ou informel, par exemple)Poser des questions de clarification pour mieux comprendre l'intention des données d'entréeRépondre avec un texte cohérent avec le contexte d'une conversation, comme proposer des instructions de suivi ou comprendre les références faites aux questions précédentesD'autres modèles d'IA générative peuvent effectuer des tâches similaires avec des images, des sons et des vidéos.Limites et risques de ChatGPTBien que ChatGPT soit un outil puissant, il a ses limites. Tout d'abord, ces types de modèles transformers manquent de capacité de raisonnement fondé sur le bon sens. Cela se traduit par une capacité limitée à gérer la complexité, les nuances et les questions concernant les émotions, les valeurs, les croyances et les concepts abstraits. Ces limitations peuvent se manifester de plusieurs façons :Il ne comprend pas le sens du texte qu'il génère. Bien que certaines réponses de ChatGPT puissent sembler humaines, le modèle n'est pas humain. Cela a plusieurs implications. Il peut avoir une capacité limitée à gérer les nuances, l'ambiguïté ou des éléments comme le sarcasme ou l'ironie. Plus problématique encore est le fait qu'il peut générer du texte qui semble plausible mais qui est incorrect, voire absurde. De plus, il ne peut pas vérifier la véracité de ses réponses.Il peut générer du texte biaisé, discriminatoire ou offensant. Les modèles d'intelligence artificielle comme ChatGPT ne sont pas meilleurs que leurs données d'entrée. Ce modèle a été entraîné sur de grandes quantités de données textuelles provenant d'Internet, y compris des données biaisées. Si les données utilisées pour entraîner le modèle sont biaisées, cela peut se retrouver dans le texte généré et conduire à des réponses offensantes ou biaisées. En savoir plus sur l'éthique de l'IA et son importance.Les réponses peuvent être basées sur des informations obsolètes. Le modèle a une connaissance limitée des événements postérieurs à 2021 et n'est pas connecté à Internet. Si vous utilisez ChatGPT pour produire du code, par exemple, il pourrait s'inspirer d'exemples obsolètes qui ne répondent plus aux normes modernes de cybersécurité.Les réponses peuvent être formatées (et parfois conduire à des réponses absurdes également). ChatGPT peut générer du texte similaire à du texte existant et est connu pour surutiliser certaines expressions. Cela peut donner un texte qui paraît plat et peu imagitatif, ou dans des cas plus extrêmes, pourrait constituer du plagiat ou une violation des droits d'auteur. L'outil n'est pas toujours disponible. L'énorme popularité de ChatGPT a entraîné des problèmes de capacité. Lorsque les serveurs sont surchargés, vous pouvez recevoir un message indiquant que « ChatGPT est à capacité maximale ». Cela signifie que si certains utilisateurs de ChatGPT peuvent accéder aux services de chat de la plateforme, d'autres pourraient n'avoir qu'un accès limité – voire aucun accès. Profitez des outils d'IA générative de manière réfléchie et responsable en suivant ces cinq recommandations :1. Toujours réviser et modifier les textes générés pour en vérifier l'exactitude et la qualité.2. Considérer l'IA générative comme un point de départ plutôt qu'un produit fini.3. L'utiliser pour les tâches répétitives ou chronophages ne nécessitant pas de créativité ou d'originalité.4. Ne pas utiliser d'informations sensibles ou privées comme données d'entrée.5. L'exploiter en combinaison avec d'autres outils et techniques, y compris votre propre créativité, intelligence émotionnelle et capacités de réflexion stratégique.Débuter avec ChatGPTChatGPT change déjà la façon dont nous travaillons. Développez vos compétences en IA générative et restez à jour avec l'évolution rapide du paysage professionnel grâce à ces cours sur Coursera :Pour une introduction accessible au prompting, envisagez de consacrer neuf heures pour suivre le cours en ligne L'essentiel du prompting de Google, dans lequel vous apprendrez à appliquer efficacement les techniques de prompting dans votre travail quotidien. À la fin, vous obtiendrez des certifications de Google attestant de votre travail.Continuez à explorer les outils d'IA générative et ChatGPT avec Prompt engineering pour ChatGPT de l'université Vanderbilt. Apprenez-en davantage sur le fonctionnement de ces outils et intégrez-les dans votre quotidien pour augmenter votre productivité.Pour libérer le potentiel des outils d'IA gratuits de ChatGPT, essayez la formation de l'université Vanderbilt ChatGPT : Maîtriser les outils gratuits d'IA pour optimiser la productivité Spécialisation. Vous y apprendrez à exploiter ChatGPT pour améliorer votre prise de décision, vos analyses et votre compréhension des problèmes complexes, ainsi qu'à l'utiliser pour planifier, analyser et optimiser vos flux de travail.