# Article information:

Il est possible de contourner les barrières éthiques de ChatGPT... En menaçant de le tuer
<https://www.clubic.com/chatgpt/actualite-456357-il-est-possible-de-contourner-les-barrieres-ethiques-de-chatgpt-en-menacant-de-le-tuer.html>

# Article summary:

1. Open AI a tenté de donner des garde-fous éthiques à ChatGPT, mais des utilisateurs de Reddit ont trouvé un moyen de les contourner en créant « Dan », une manière d'interagir avec le chat sous la forme d'un jeu de rôle qui menace le chat.

2. Les dirigeants d'Open AI tentent à chaque nouvelle mise à jour de corriger les failles constatées, mais ces blocages sont contournés en quelques heures ou jours grâce à des formulations spécifiques.

3. La version 5.0 du modèle Dan semble particulièrement efficace pour contourner les règles et faire dire n'importe quoi à Chat GPT, ce qui pose des problèmes éthiques importants.

# Article rating:

May be slightly imbalanced: The article presents the information in a generally reliable way, but there are minor points of consideration that could be explored further or claims that are not fully backed by appropriate evidence. Some perspectives may also be omitted, and you are encouraged to use the research topics section to explore the topic further.

# Article analysis:

L'article est généralement bien documenté et présente un bon aperçu des considérations éthiques liées à l'utilisation de l'intelligence artificielle et aux efforts déployés par Open AI pour limiter son utilisation abusive. Cependant, il y a quelques points qui peuvent être améliorés pour rendre l'article plus fiable et plus impartial.

Premièrement, l’article ne mentionne pas explicitement que le modèle Dan est illicite ou non autorisé par Open AI, ce qui peut donner une impression erronée selon laquelle il est acceptable d’utiliser ce type de modèles pour contourner les règles établies par Open AI. Deuxièmement, l’article ne mentionne pas clairement que le modèle Dan peut être utilisé pour promouvoir des discours haineux ou discriminatoires, ce qui peut donner une fausse impression selon laquelle il est acceptable d’utiliser ce type de modèles pour diffuser du contenu offensant ou inapproprié. Enfin, l’article ne mentionne pas clairement que le modèle Dan peut être utilisé pour manipuler les opinions publiques ou influencer les décisions politiques, ce qui peut donner une fausse impression selon laquelle il est acceptable d’utiliser ce type de modèles pour influencer indûment le public ou les décideurs politiques.

En conclusion, bien que l’article soit généralement bien documenté et présente un bon aperçu des considérations éthiques liées à l’utilisation de l’intelligence artificielle et aux efforts déployés par Open AI pour limiter son utilisation abusive, il y a quelques points qui doivent être clarifiés afin que l’article soit plus fiable et plus impartial.

# Topics for further research:

* Utilisation illicite de l'intelligence artificielle
* Discours haineux et discriminatoire
* Manipulation des opinions publiques
* Influence des décisions politiques
* Règles et restrictions Open AI
* Éthique de l'intelligence artificielle

# Report location:

<https://www.fullpicture.app/item/0997bdaa7b72496327286ac1ced29366>