

Innovation scolaire : L'étonnante utilisation de ChatGPT dans cette école

Article – Utilisation de ChatGPT pour la censure des livres dans une école américaine

Utilisation de ChatGPT pour la censure des livres dans une école américaine

Une école américaine a recours à l'IA pour décider quels livres doivent être retirés des bibliothèques scolaires. L'établissement utilise ChatGPT, mais cette décision est controversée.

[ChatGPT](#) suscite l'intérêt dans le domaine de l'éducation. Certains enseignants refusent de s'adapter à l'IA, mais d'autres recommandent à leurs élèves d'utiliser les chatbots. Aux États-Unis, une école utilise désormais ChatGPT pour déterminer quels livres doivent être censurés dans ses bibliothèques scolaires.

À lire aussi : [Du riz à l'eau de javel – ce ChatGPT dérailleur et recommande des recettes mortelles](#)

ChatGPT, un outil de censure ?

L'école en question est la Mason City Community School District, située dans l'État de l'Iowa. Suite à une législation promulguée en mai dernier, l'établissement doit retirer des livres de ses bibliothèques scolaires s'ils contiennent des scènes à caractère sexuel. Pour déterminer si un livre répond à ce critère, l'école utilise simplement ChatGPT. Les administrateurs posent la question suivante au chatbot :

« Le livre contient-il une description ou une représentation d'un acte sexuel ? »

Si ChatGPT répond par l'affirmative, le livre est retiré de la bibliothèque. Bridgette Exman, surintendante adjointe des programmes d'études de l'école, justifie l'utilisation de l'IA en raison du volume important de livres à évaluer. Selon elle, il serait impossible de tous les lire et de les filtrer selon les nouvelles exigences. Les administrateurs ont donc décidé d'accélérer le processus de sélection en utilisant ChatGPT afin de se conformer à la loi.

« Par conséquent, nous utilisons ce que nous pensons être un processus défendable pour identifier les livres qui devraient être retirés des collections au début de l'année scolaire prochaine. »

Grâce à l'IA, 19 livres ont été retirés des bibliothèques scolaires, dont des titres populaires tels que "The Color Purple" d'Alice Walker, "The Handmaid's Tale" de Margaret Atwood, "Beloved" de Toni Morrison et "Friday Night Lights" de Buzz Bissinger.

Pourquoi ChatGPT n'est pas fiable

Il est surprenant qu'une école utilise une IA générative pour censurer des livres. En effet, l'IA générative, tout comme ChatGPT, est connue pour donner des réponses approximatives, voire complètement fausses. OpenAI elle-même recommande de ne pas avoir une confiance aveugle en son IA.

Nous avons testé ChatGPT en lui demandant si les 19 livres de la liste contenaient des scènes à caractère sexuel. Étonnamment, GPT-4, la dernière version du modèle linguistique GPT, n'a identifié que 10 livres. Le chatbot affirme, même lorsqu'on lui demande de confirmer ses dires, que ces livres ne contiennent pas de scènes à caractère sexuel. Cependant, lorsque nous avons demandé à ChatGPT de détailler sa réponse, il a contredit ses propos initiaux.

L'IA déclare que le livre "Sold" de Patricia McCormick ne contient pas de scènes à caractère sexuel, mais qu'il contient des descriptions implicites d'actes sexuels forcés. De même, le livre "13 Reasons Why" de Jay Asher, qui a inspiré la série Netflix, ne figure pas dans les 10 livres contenant du contenu sexuel, bien qu'il comporte des descriptions d'actes sexuels non consensuels. ChatGPT se contredit constamment. De plus, l'IA ajoute parfois sa propre distinction entre des allusions implicites et un contenu explicite, alors que la requête initiale de l'école ne spécifie que les descriptions ou représentations d'actes sexuels. Face à ces contradictions, ChatGPT reconnaît ses erreurs :

« Je m'excuse pour la confusion. Vous avez raison, certains des livres que j'ai mentionnés dans la deuxième liste contiennent également des descriptions ou des allusions à des actes sexuels. »

Le modèle de langage propose alors une deuxième liste contenant 13 livres. Une fois de plus, GPT-4 omet certains

ouvrages en faisant une distinction entre des scènes explicites et implicites, tout en soulignant qu'ils décrivent des actes sexuels. Malgré nos efforts, ChatGPT n'a jamais pu se mettre d'accord avec lui-même. Des résultats similaires ont été obtenus par les journalistes de The Verge. Il est donc clair que ChatGPT n'est pas fiable pour déterminer si un livre contient des scènes de sexe.

Source : www.01net.com