

OpenAI confronté à des poursuites judiciaires pour les hallucinations de ChatGPT

Une fausse information relayée par ChatGPT

OpenAI, une entreprise spécialisée dans l'intelligence artificielle et notamment dans la création de chatbots, est visée par une plainte de Mark Walters, animateur radio américain pro-armes et PDG de CCW Broadcast Media. La plainte fait suite à une réponse de ChatGPT, l'un des outils d'OpenAI, à une demande du rédacteur en chef d'un site spécialisé dans les armes à feu, Fred Riehl. Ce dernier souhaitait résumer une affaire concernant le Procureur général de Washington, Bob Ferguson.

La réponse de ChatGPT contenait une information fausse selon laquelle Walters aurait détourné des fonds de l'association SAF (The Second Amendment Foundation), une association de défense du port d'armes, et qu'il serait trésorier et directeur financier de celle-ci. Pourtant, Walters n'a aucun lien avec SAF.

Une plainte pour diffamation

Si Fred Riehl n'a pas repris l'information produite par ChatGPT, Mark Walters estime que cette information fausse porte atteinte à sa réputation et a porté plainte pour diffamation, demandant des dommages et intérêts qui seront déterminés lors du procès. Cette plainte pourrait créer un

précédent et encourager d'autres personnalités à porter plainte contre ChatGPT pour des informations fausses publiées à leur sujet.

Ce qu'en dit la justice

Selon Eugene Volokh, professeur à la faculté de droit de l'Université de Californie à Los Angeles, l'erreur produite par ChatGPT pourrait constituer une diffamation si le plaignant peut prouver qu'il a subi un préjudice, tel qu'une perte d'emploi ou de revenu, à cause de cette information fausse. Cependant, il reste à déterminer si ChatGPT peut être considéré comme responsable de la diffusion de cette information puisque c'est un outil d'intelligence artificielle et que les informations qu'il produit ne sont pas validées par des humains avant d'être publiées.

Source : www.lemondeinformatique.fr

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