

Claude brille, ChatGPT dans l'ombre : le duel des chatbots !

Les Chatbots en Suisse : Réalité et Défis

Dans un monde de plus en plus numérique, les chatbots intelligents font leur apparition dans divers domaines, de la santé à la défense des consommateurs. Cependant, un récent test de la Haute école spécialisée du Nord-Ouest de la Suisse (FHNW), diffusé dans l'émission Kassensturz de SRF, révèle que même des leaders comme ChatGPT éprouvent des difficultés lorsqu'il s'agit de répondre à des questions typiquement suisses. Contrairement à Claude, un autre chatbot, qui s'avère plus performant, ChatGPT semble moins convaincant.

Des Réponses Inattendues

Lors du test, une situation humoristique mais révélatrice a été mise en avant. Quand on interrogeait deux chatbots sur ce qu'on obtient dans un restaurant suisse alémanique après avoir commandé une "Strange", les réponses étaient confuses. Les chatbots affirmaient qu'il s'agissait d'une pâtisserie allongée, alors qu'en réalité, une "Strange" se réfère à une chope, souvent utilisée pour la bière. Ce type d'erreur, connu sous le nom d'"hallucination", est une faiblesse commune parmi de nombreux modèles d'intelligence artificielle.

Le Test des Chatbots

Pour évaluer les performances des chatbots, les chercheurs de la FHNW ont soumis dix modèles largement utilisés à plus de 300 questions. Ces questions touchaient divers domaines, y compris la culture, le droit, et particulièrement des questions suisses. À leur grande surprise, la plupart des chatbots ont montré des lacunes significatives dans leur capacité à répondre correctement aux questions liées à la Suisse.

ChatGPT : Suffisamment Efficace, Mais Pas Plus

Bien que ChatGPT soit un des chatbots les plus utilisés avec environ 800 millions d'utilisateurs interagissant chaque semaine, les résultats indiquent que ses performances ne sont que "suffisantes". Selon Simon Felix, expert en intelligence artificielle, ChatGPT ne produit ni de mauvaises ni d'excellentes réponses et se heurte à des difficultés avec des questions typiquement suisses.

Claude : La Révélation

En revanche, le chatbot Claude, développé par Anthropic, a dominé le test. Ce modèle a fourni les réponses les plus pertinentes et concises, rendant compte d'une approche particulièrement efficace. Simon Felix note que Claude va droit au but, sans omettre des détails cruciaux. En outre, Copilot de Microsoft et Gemini de Google ont également montré de solides résultats, bien que les réponses de Copilot soient influencées par le même modèle que ChatGPT, illustrant comment les directives des développeurs peuvent améliorer les résultats.

Une Utilisation Précautionneuse Nécessaire

Alors que l'utilisation des chatbots IA se généralise, il est essentiel d'aborder les sujets sensibles avec prudence. Les modèles de langage tels que ChatGPT devraient être réservés aux cas où les conséquences d'une mauvaise réponse ne sont pas graves. Comme le souligne Simon Felix, une erreur dans des questions concernant la médecine ou le droit pourrait avoir des implications dramatiques.

Enfin, il est impératif de protéger les données sensibles lors de l'utilisation des chatbots. Les informations privées n'ont pas leur place dans ces systèmes, et une gestion appropriée des données est incontournable pour éviter les fuites.



Article inspiré de l'œuvre de Pirmin Roos (SRF), adapté par Julien Furrer (RTS).

Source : www.rts.ch

→ Accéder à CHAT GPT en cliquant dessus