Cybrarians Journal
E-ISSN 1687-2215

العدد 76، أغسطس 2025
DOI: 10.70000/cj.2025.76.642

تشات جي بي تي ومستخدمو المكتبات: مخاطر هلوسة الذكاء الاصطناعي والمعلومات المضللة

بحوث – الملخص العربي

تاريخ الاستلام: 15 مارس 2025
تاريخ القبول: 15 يوليو 2025
تاريخ النشر: 31 أغسطس 2025

بولاجي أولادوكون
الجامعة الفيدرالية للتكنولوجيا، إيكوت أباسي، نيجيريا
Bolaji.oladokun@yahoo.com

فيفيان أولوتشي إيمانويل
جامعة ولاية ريفرز، بورت هاركورت، نيجيريا
Emmanuel.vivien@ust.edu.ng

أونوم كوين أوساجي
جامعة جنوب أفريقيا، جنوب أفريقيا
Onomeosagie@gmail.com

أديفونكي أولانيك ألابي
جامعة لاجوس، ولاية لاجوس، نيجيريا
oladesh@yahoo.com

المستخلص

يكشف التحليل أن الاعتماد المتزايد على أدوات الذكاء الاصطناعي مثل ChatGPT من قبل مستخدمي المكتبات يطرح مخاطر كبيرة تهدد نزاهة البحث وجودته. في حين تقدم هذه الأدوات مزايا ملحوظة من حيث الكفاءة وسرعة الوصول إلى المعلومات، إلا أنها تنطوي على عيوب جوهرية، أبرزها ظاهرة "هلوسات الذكاء الاصطناعي"، وهي ميل النموذج إلى توليد معلومات تبدو معقولة ولكنها غير صحيحة أو مختلقة بالكامل.

تشمل المخاطر الرئيسية نشر المعلومات المضللة، وتآكل مهارات التفكير النقدي لدى المستخدمين، وتفاقم المخاوف الأخلاقية المتعلقة بالتحيز في البيانات وحقوق الملكية الفكرية. ونتيجة لذلك، قد يؤدي الاعتماد غير الخاضع للرقابة على ChatGPT إلى تدهور جودة المخرجات الأكاديمية وتقويض دور أمناء المكتبات كمرشدين موثوقين في عملية البحث.

يؤكد البحث على الدور المحوري الذي يجب أن تلعبه المكتبات وأمناؤها في مواجهة هذه التحديات. يتضمن ذلك تثقيف المستخدمين بشكل استباقي حول حدود الذكاء الاصطناعي، وتعزيز ثقافة التقييم النقدي للمحتوى الذي يتم إنشاؤه بواسطة الذكاء الاصطناعي، والتأكيد على ضرورة التحقق من المعلومات عبر مصادر موثوقة. إن تبني استراتيجيات التخفيف، مثل تحسين بيانات التدريب والرقابة البشرية، أمر حاسم لضمان الاستخدام المسؤول لهذه التقنيات القوية.

الكلمات المفتاحية

الذكاء الاصطناعي، تشات جي بي تي، الهلوسة، المعلومات المضللة

نظرة عامة على ChatGPT وتقنيته الأساسية

ChatGPT، الذي طورته شركةOpenAI ، يمثل تقدمًا كبيرًا في مجال معالجة اللغات الطبيعية. يعتمد النموذج على بنية GPT-4 (المحول التوليدي المدرب مسبقًا 4)، وهو نموذج لغوي متطور يمثل الجيل الرابع في سلسلة النماذج التي صممتها.OpenAI

تتضمن عملية تطوير GPT-4 مرحلتين رئيسيتين:

1. **التدريب المسبق :(Pre-training)** يتم تدريب النموذج على مجموعة بيانات نصية ضخمة ومتنوعة من الإنترنت، مما يسمح له بتعلم الأنماط والهياكل والفروق الدقيقة في اللغة البشرية.

2. **الضبط الدقيق :(Fine-tuning)** يتم تحسين النموذج بشكل إضافي باستخدام مجموعة بيانات أضيق مع إشراف بشري لتعزيز دقته وملاءمته.

تعتمد تقنية GPT-4 على بنية "المحول(Transformer) "، التي تتكون من طبقات متعددة من "آليات الانتباه .(Attention Mechanisms) "تمكّن هذه الآليات النموذج من تقييم أهمية الكلمات المختلفة في

الجملة، مما يؤدي إلى فهم السياق بشكل أكثر فعالية مقارنة بالنماذج السابقة. وقد ساهمت هذه الابتكارات في قدرة GPT-4 على إنتاج نصوص شبيهة بالنصوص البشرية يصعب تمييزها أحيانًا عن تلك التي يكتبها الإنسان.

ظاهرة هلوسات الذكاء الاصطناعي

على الرغم من التقدم الكبير في الذكاء الاصطناعي، فإن أنظمة مثل ChatGPT يمكن أن تظهر ظاهرة تُعرف باسم "الهلوسات". تشير هلوسات الذكاء الاصطناعي إلى الحالات التي يولد فيها الذكاء الاصطناعي معلومات أو استجابات تبدو معقولة ومنطقية ولكنها غير صحيحة من الناحية الواقعية أو لا معنى لها. تتميز هذه الهلوسات بقابليتها الخادعة للتصديق، مما يجعل من الصعب اكتشافها وتصحيحها.

أمثلة على الهلوسات:

- **الحقائق التاريخية المزيفة :**قد يقدم النموذج تفاصيل واثقة ومختلقة بالكامل حول حياة وإنجازات شخصية تاريخية.
- **البيانات العلمية الوهمية :**يمكن أن يولد النموذج بيانات علمية أو نصائح طبية تبدو معقولة ولكنها غير صحيحة، مما قد يضلل المستخدمين.

أسباب هلوسات الذكاء الاصطناعي:

- **طبيعة بيانات التدريب :**يتم تدريب النماذج على مجموعات بيانات ضخمة من الإنترنت تحتوي على معلومات دقيقة وغير دقيقة على حد سواء. يتعلم الذكاء الاصطناعي الأنماط ولكنه لا يطور فهمًا حقيقيًا للمحتوى.
- **الطبيعة الاحتمالية للنماذج :**تولد هذه النماذج الاستجابات بناءً على احتمالية تسلسل الكلمات، مما قد يؤدي إلى إنشاء معلومات معقولة ولكنها غير صحيحة، خاصة عند مواجهة فجوات في بيانات التدريب.
- **تعقيد اللغة البشرية :**تكافح نماذج الذكاء الاصطناعي لفهم الفروق الدقيقة والسياق والمعرفة الضمنية في اللغة البشرية بشكل كامل، مما قد يؤدي إلى استجابات غير دقيقة.
- **التحيزات المتأصلة :**إذا كانت بيانات التدريب تحتوي على تحيزات أو معلومات مضللة، فيمكن للنموذج نشرها وتضخيمها في استجاباته.

المخاطر المتصورة للاعتماد على ChatGPT في البحث وإنشاء المحتوى

يرافق الاعتماد المتزايد على ChatGPT مخاطر كبيرة تؤثر على جودة ونزاهة البحث وإنشاء المحتوى.

DOI: 10.70000/cj.2025.76.642

فئة المخاطر	الوصف التفصيلي
نشر المعلومات المضللة	الخطر الأكثر إلحاحًا هو قدرة ChatGPT على توليد استجابات غير صحيحة من الناحية الواقعية. في الأبحاث الأكاديمية والمهنية، حيث تكون الدقة أمرًا بالغ الأهمية، يمكن أن يؤدي ذلك إلى نشر معلومات خاطئة وتقويض مصداقية المحتوى.
الافتقار إلى التحقق من المصادر	يفتقر النموذج إلى القدرة على التحقق من صحة ومصداقية المصادر التي يشير إليها. وبالتالي، قد يعتمد المحتوى الذي يتم إنشاؤه على معلومات غير موثوقة أو متحيزة، مما قد يؤدي إلى إدخال أخطاء في عمل الباحثين.
المخاوف الأخلاقية	يثير استخدام ChatGPT قضايا أخلاقية متعددة، بما في ذلك إمكانية إدامة التحيزات الموجودة في بيانات التدريب (مثل التحيزات المتعلقة بالعرق أو الجنس) وإثارة تساؤلات حول التأليف وحقوق الملكية الفكرية.
تآكل التفكير النقدي	قد يؤدي الاعتماد المفرط على الذكاء الاصطناعي إلى تقليل انخراط الباحثين في عمليات التقييم والتركيب والتفكير الأصيل. مع مرور الوقت، يمكن أن يؤدي تآكل هذه المهارات إلى انخفاض جودة العمل الأكاديمي.
التبعية والإفراط في الاعتماد	قد يصبح المستخدمون أقل ميلًا للانخراط في أساليب البحث التقليدية وممارسات الكتابة، مثل المراجعة الشاملة للأدبيات وتحليل البيانات. على المدى الطويل، يمكن أن يؤثر هذا التحول على عمق ودقة البحث.

الآثار المترتبة على استخدام ChatGPT من قبل مستخدمي المكتبات

إن دمج ChatGPT في خدمات المكتبات له آثار سلبية كبيرة على جودة البحث والنزاهة الأكاديمية.

- **تدهور جودة البحث :**نظرًا لأن استجابات ChatGPT ليست دقيقة أو موثوقة دائمًا، فإن استخدامها في الأوساط الأكاديمية يمكن أن يؤدي إلى نتائج بحثية معيبة، ونشر معلومات خاطئة، وانخفاض عام في جودة العمل العلمي.

•**تقويض نزاهة البحث :**إن سهولة استخدام ChatGPT تشجع على عقلية "الطريق المختصر"، حيث يقبل المستخدمون المحتوى الذي يتم إنشاؤه دون التحقق الدقيق. هذا التحول يقوض المهارات البحثية الأساسية مثل التفكير النقدي والتقييم الدقيق للمصادر.

•**تآكل مهارات التفكير النقدي :**قد يؤدي الاعتماد على الاستجابات التي يولدها الذكاء الاصطناعي إلى تقليل ميل المستخدمين إلى الانخراط في التحليل المستقل والتقييم النقدي للمعلومات، مما يؤدي إلى فهم سطحي ومخرجات بحثية ضعيفة.

•**تضاؤل دور أمناء المكتبات :**مع تحول المستخدمين بشكل متزايد إلى ChatGPT للحصول على إجابات فورية، قد ينخفض الطلب على خبرة أمناء المكتبات. هذا التحول يمكن أن يقلل من الدعم الشخصي والتوجيه الذي يقدمونه، مما يؤدي إلى بيئة بحثية أقل قوة.

•**خلق التبعية :**يفتقر ChatGPT إلى الفهم الدقيق والوعي السياقي الذي يقدمه الخبراء البشريون. يمكن أن يخلق الاعتماد المفرط على الذكاء الاصطناعي تبعية تحد من قدرة المستخدمين على إجراء بحث مستقل وحل المشكلات، خاصة عندما تكون الاستجابات التي يولدها الذكاء الاصطناعي غير صحيحة.

استراتيجيات التخفيف لمنع هلوساتChatGPT

لمواجهة تحديات هلوسات الذكاء الاصطناعي، يقترح البحث عدة استراتيجيات تخفيف فعالة:

1. **تحسين بيانات التدريب :**تتمثل إحدى أكثر الاستراتيجيات فعالية في تحسين جودة بيانات التدريب عن طريق تنظيم مجموعات بيانات من مصادر حسنة السمعة وتحديثها باستمرار. إن تدريب الذكاء الاصطناعي على بيانات عالية الجودة يمكن أن يقلل بشكل كبير من احتمالية توليد الهلوسات.

2. **الضبط الدقيق :(Fine-tuning)** يتضمن ذلك تدريب النموذج المدرب مسبقًا على مجموعة بيانات أضيق وأكثر تخصصًا. على سبيل المثال، سيكون النموذج الذي تم ضبطه بدقة باستخدام الأدبيات الطبية أكثر موثوقية في الإجابة على الأسئلة المتعلقة بالصحة.

3. **آليات التحقق من الحقائق :**من الضروري دمج آليات قوية للتحقق من الحقائق في عملية توليد الاستجابات. تتضمن هذه الآليات التحقق المتقاطع من المحتوى الذي يتم إنشاؤه بواسطة الذكاء الاصطناعي مع قواعد بيانات ومصادر موثوقة في الوقت الفعلي.

DOI: 10.70000/cj.2025.76.642

4. **دمج ملاحظات المستخدمين :**يعد دمج ملاحظات المستخدمين بشكل فعال استراتيجية حاسمة أخرى. يمكن للمستخدمين الذين يواجهون استجابات غير صحيحة تقديم ملاحظات يمكن تحليلها واستخدامها لتحسين النموذج.

5. **تحديد نطاق التطبيقات :**يمكن أن يساعد تحديد نطاق تطبيقات الذكاء الاصطناعي ووضع حدود واضحة لاستخدامه في منع الهلوسات. يقلل هذا من احتمالية خوض النموذج في مجالات يفتقر فيها إلى بيانات تدريب كافية أو فهم سياقي.

الخاتمة والتوصيات

يخلص البحث إلى أن الاعتماد على ChatGPT في البحث وإنشاء المحتوى من قبل مستخدمي المكتبات يمثل مخاطر كبيرة، خاصة بسبب احتمالية حدوث هلوسات الذكاء الاصطناعي. يمكن أن تؤدي هذه الهلوسات إلى نشر معلومات مضللة، وتقويض نزاهة العمل العلمي، وتآكل مهارات التفكير النقدي.

التوصيات للسياسات والممارسات:

•**على مستوى السياسات :**يجب على المكتبات وضع مبادئ توجيهية للاستخدام الأخلاقي لأدوات الذكاء الاصطناعي. يجب على صانعي السياسات إعطاء الأولوية للمبادرات التي تدمج محو الأمية المعلوماتية النقدية في خدمات المكتبات وفرض أطر عمل للتحقق من الحقائق.

•**على المستوى العملي :**يجب على المكتبات أن تتخذ دورًا استباقيًا في تثقيف المستخدمين حول مخاطر المعلومات المضللة. يجب تزويد أمناء المكتبات بالمهارات اللازمة لتوجيه المستخدمين في التقييم النقدي للاستجابات التي يولدها الذكاء الاصطناعي والتحقق منها. يجب على المكتبات الاستثمار في تدريب الموظفين وورش العمل التي تركز على محو الأمية في مجال الذكاء الاصطناعي.

من خلال تنفيذ هذه الاستراتيجيات، يمكن للمكتبات تعزيز فوائد الذكاء الاصطناعي مع حماية جودة وموثوقية المعلومات، ودعم المساعي العلمية والبحثية لمستخدميها في نهاية المطاف.

DOI: 10.70000/cj.2025.76.642