خطورة الذكاء الاصطناعي التوليدي

جمال مراد قيس

2025-08-17

يشهد العالم ثورة معرفية وإبداعية بفضل تقنيات الذكاء الاصطناعي التوليدي، التي أصبحت قادرة على إنتاج نصوص، صور، مقاطع صوتية ومرئية، بل وحتى أكواد برمجية، بدرجة من الإتقان كانت قبل سنوات قليلة حكرًا على البشر. ومع هذا التقدم، ظهرت تحديات عميقة تتعلق بالجانب الأخلاقي والتقني لهذا المحتوى المولّد، ما استدعى إجراء أبحاث أكاديمية ومهنية لتقييم المخاطر ووضع الضوابط.

من الناحية التقنية، تركز الأبحاث على فهم الآليات الداخلية لنماذج الذكاء الاصطناعي التوليدي مثل النماذج اللغوية الضخمة (LLMs) والشبكات التوليدية الخصامية (GANs)، بهدف تحسين دقتها، تقليل الأخطاء، ومعالجة مشكلة الانحيازات المدمجة في البيانات. إن هذه النماذج تعتمد على كميات هائلة من البيانات المأخوذة من مصادر متنوعة، مما يعني أن أي تحيز أو خطأ موجود في البيانات الأصلية يمكن أن ينتقل إلى المحتوى الناتج. ولهذا السبب، يجتهد الباحثون في تصميم خوارزميات أكثر شفافية وقابلية للتفسير، تمكّن المستخدمين والمطورين من فهم كيفية اتخاذ النموذج لقراراته الإبداعية.

أما من الجانب الأخلاقي، فإن الإشكالية الأكثر بروزًا تتمثل في مسألة الأصالة والمصداقية، فالمحتوى المولّد يمكن أن يكون واقعيًا لدرجة تصعب معها تمييزه عن المحتوى البشري، مما يفتح الباب أمام استخدامه في التضليل الإعلامي أو نشر الأخبار الزائفة أو إنشاء مواد مضللة تستهدف الرأي العام. كما يشير الباحثون إلى ضرورة وضع سياسات واضحة لتحديد ملكية حقوق النشر للمحتوى المولّد، خصوصًا عندما يتداخل الإبداع البشري مع المخرجات الآلية في مشروع واحد.

وتثير الأبحاث أيضًا مسألة المسؤولية القانونية، إذ لا يزال الجدل قائمًا حول من يتحمل تبعات المحتوى المولّد إذا تضمن معلومات مضللة أو مسيئة: هل هو المطور، أم المستخدم، أم الشركة المالكة للتقنية؟ كما يطرح هذا الجدل

تساؤلات حول المعايير الدولية التي ينبغي وضعها لتوحيد السياسات عبر الدول، بما يضمن حماية المستخدمين وحقوق الأطراف المختلفة.

إلى جانب ذلك، هناك اهتمام متزايد بمفهوم "السلامة" في النماذج التوليدية، أي ضمان ألا تنتج هذه النماذج محتوى ضارًا أو غير قانوني حتى عند الاستفزاز أو التلاعب بالمُدخلات. بعض الأبحاث تقترح دمج فلاتر أمان متعددة الطبقات، وأخرى تركز على تدريب النماذج على الاستجابة ضمن حدود أخلاقية واضحة. ويؤكد الباحثون أن التوازن بين حرية الإبداع والقيود الأخلاقية هو معركة مستمرة، لأن فرض ضوابط صارمة قد يحد من الابتكار، بينما ترك المجال مفتوحًا بلا ضوابط قد يؤدي إلى أضرار جسيمة.

المحاور العملية المرتبطة بالأبحاث الأخلاقية والتقنية للمحتوى المولّد بالذكاء الاصطناعي:

- **دقة المخرجات**: ضرورة التحقق من صحة المعلومات التي ينتجها الذكاء الاصطناعي لتفادي الأخطاء والمغالطات. - **الشفافية**: تمكين المستخدمين من معرفة ما إذا كان المحتوى مولّدًا آليًا، وتوضيح آلية إنتاجه. - **التحيزات المدمجة**: معالجة الانحيازات في البيانات التي قد تنعكس على المحتوى المولّد وتؤدي إلى نتائج غير عادلة أو مضللة. - **حماية الملكية الفكرية**: تحديد حقوق النشر للمحتوى الناتج عن الدمج بين إبداع بشري ومخرجات آلية. - **الأمن والسلامة الرقمية**: منع توليد محتوى ضار أو غير قانوني عبر فلاتر ذكية ومتعددة الطبقات. - **المساءلة القانونية**: وضع أطر واضحة لتحديد المسؤولية عند حدوث انتهاكات أو أضرار بسبب المحتوى المولّد.

لا تنحصر التحديات التقنية في مرحلة تدريب النماذج التوليدية فحسب، بل تمتد إلى مرحلة ما بعد الإطلاق، حيث يصبح من الضروري مراقبة المخرجات بشكل دوري لضمان توافقها مع الضوابط الأخلاقية والمعايير التقنية المعتمدة. ولتحقيق ذلك، يقترح بعض الخبراء تأسيس جهات رقابية متخصصة تتولى مراجعة وفحص المحتوى المولّد، على غرار اللجان الأخلاقية في مجال البحوث العلمية أو الهيئات المعنية بمراجعة الدراسات الطبية، بما يضمن بيئة إنتاج رقمية أكثر أمانًا ومسؤولية.

توضح الأبحاث أن الحلول الفعّالة تحتاج إلى تعاون متعدد الأطراف، يجمع المطورين، الأكاديميين، المشرعين، والمجتمع المدني، للوصول إلى أطر تنظيمية ديناميكية قادرة على مواكبة التطورات السريعة في هذا المجال. ومن خلال هذا التعاون، يمكن صياغة معايير تقنية وأخلاقية مرنة، تسمح بالاستفادة الكاملة من قدرات الذكاء الاصطناعي التوليدي، مع تقليل مخاطره على الأفراد والمجتمعات.

المصادر

- ,Floridi, L., & Chiriatti, M. (2020). GPT-3: Its Nature, Scope, Limits .and Consequences .Minds and Machines ,30(4) ,681–694.
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI .ethics guidelines .Nature Machine Intelligence ,1(9) ,389–399.
- Weidinger, L., Mellor, J., Rauh, M., et al. (2021). Ethical and social .risks of harm from Language Models .arXiv preprint arXiv: .2112.04359
- Weapons of Math Destruction: How Big Data .O'Neil, C. (2016) .Increases Inequality and Threatens Democracy .Crown Publishing .Group
- .Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S .On the Dangers of Stochastic Parrots: Can Language .(2021) .Proceedings of the 2021 ACM Conference on ?Models Be Too Big .Fairness, Accountability, and Transparency

تواصل مع الكاتب: mohamedmouradgamal@gmail.com