إنستجرام يدعو منصات التواصل لتعزيز الشفافية بشأن المحتوى المُنتج بالذكاء الاصطناعى

أثار آدم موسيري، رئيس منصة إنستجرام، قضية تنامي استخدام الذكاء الاصطناعي التوليدي في إنشاء منشورات وصور على وسائل التواصل الاجتماعي، داعيًا إلى ضرورة توفير مزيد من الشفافية حول هذا النوع من المحتوى.

وفي سلسلة منشورات على منصة Threads التابعة لشركة ميتا، أكد موسيري أن الخط الفاصل بين المحتوى الواقعي والمحتوى المُنتج بالذكاء الاصطناعي أصبح غامضًا بشكل كبير، مما يتطلب تقديم سياق إضافي لمساعدة المستخدمين على التحقق من مصداقية المحتوى.

وقال موسيري: "سواء كنت متفائلًا أو متحفظًا بشأن هذه التكنولوجيا، فإن الذكاء الاصطناعي التوليدي ينتج محتوى يصعب تمييزه عن الواقع، ويتحسن بسرعة فائقة."

وأشار إلى أن منصات الإنترنت تبذل جهودًا لوضع علامات تعريفية على المحتوى المُنتج بالذكاء الاصطناعي "قدر المستطاع"، لكن بعض المحتوى قد ينجو من هذه التصنيفات، مما قد يؤدي إلى تضليل المستخدمين، لا سيما أن "ليس كل التزييفات سيتم إنتاجها باستخدام الذكاء الاصطناعي".

وأكد موسيري على أهمية تقديم معلومات إضافية حول "هوية ناشر المحتوى"، بما يتيح للمستخدمين تقييم مدى مصداقيته، وتأتي هذه الدعوة في ظل الانتشار الكبير للمحتوى المُنتج بالذكاء الاصطناعي على منصات التواصل الاجتماعي، حيث تتيح خدمات مثل Llama التابعة لميتا، وChatGPT وE-DALL من OpenAI، وGrok-2 من X للمستخدمين إنشاء صور ومنشورات فائقة الواقعية، يصعب على كثير من المستخدمين تمييزها عن الواقع.

وأشار موسيري إلى أن بعض المنصات، مثل إكس وBlueSky، تعتمد أنظمة مراقبة وتصفية لتقديم سياق إضافي حول المحتوى، فعلى سبيل المثال توفر منصة إكس ميزة "ملاحظات المجتمع" (Notes Community)، بينما تقدم BlueSky فلاتر مخصصة للمحتوى.

ومع ذلك، لا تمتلك ميتا حتى الآن نظامًا مشابهًا، ما يجعل دعوة موسيري خطوة أولى نحو تبني أدوات جديدة لتوفير سياق إضافي حول المحتوى المُنتج بالذكاء الاصطناعي.

ورغم أن ميتا لم تعلن رسميًا عن كيفية تطبيق هذه الأدوات، أشار موسيري إلى احتمالية تقديم ميزات مشابهة لتلك التي تعتمدها منصات مثل BlueSky، كما يتوقع أن تُدمج هذه الأدوات في تحديثات قادمة لتعزيز الشفافية والمصداقية على منصات ميتا، بما في ذلك ثريدز.