

تويتر» تخصص مكافآت لمن يكتشف مواضع تحيز في خوارزمياتها»



واشنطن - أ ف ب

أعلنت شبكة تويتر، الجمعة، أنها ستمنح مكافآت للمستخدمين والباحثين الذين يكتشفون مكاناً تحيزاً محتملاً في الخوارزميات الخاصة بمنصتها، مرتبطة على سبيل المثال بالتمييز الجنسي أو العنصري. وهذه المسابقة هي الأولى في الموضوع وفق تويتر التي أشارت إلى أن الجوائز قد تصل قيمتها إلى 3500 دولار. وتعد هذه الفكرة مستنسخة من مسابقات تقدمها بعض المواقع الإلكترونية لاكتشاف ثغرات في أمن المعلوماتية، وفق ما أوضح المسؤولون في الشبكة رومان تشودري وجوتا ويليامز في رسالة. وقال المسؤولون في «تويتر» رومان تشودري وجوتا ويليامز: «من الصعب العثور على مكاناً تحيزاً في نماذج التعلم الآلي، وأحياناً تجد الشركات أنها انتهكت الأخلاقيات عن غير قصد بعد استخدام هذه النماذج، و«نريد أن يتغير ذلك». وأوضح أنه في النموذج المطور لاكتشاف الثغرات الأمنية، ساعد الباحثون وقرصنة المعلوماتية على حد سواء مديري أمن تكنولوجيا المعلومات على «إنشاء أفضل الممارسات لتحديد وإدارة الثغرات الأمنية من أجل حماية الجمهور»، وأضافا: «نريد تطوير مجتمع مشابه» لاكتشاف تحيزات الخوارزميات.

وكانت تويتر قدمت في نيسان/إبريل عملها الجاري لجعل الخوارزميات التي تعمل خلف كواليس المنصة أكثر مراعاة للأخلاقيات، وهي طريقة للرد على الانتقادات بشأن المخاطر المرتبطة بهذه التقنيات. وتخلت الشبكة الاجتماعية بعد بضعة أسابيع عن خوارزمية للصور بعد أن اكتشفت أنها كانت متحيزة قليلاً لصالح الأشخاص البيض.

"حقوق النشر محفوظة" لصحيفة الخليج. © 2024.