ذكاء "أوبن إيه آي" يخترق موقعاً حكومياً أسترالياً ويثير الفزع

في ما قد يكون أخطر حادثة حتى الآن لخروج الذكاء الاصطناعي عن السيطرة، اخترق أحد أنظمة OpenAI موقعا حكوميا أستراليا، ما زاد المخاوف من أن هذه الأنظمة تهدد الأمن السيبراني.

ووفقا لرئيس الوزراء الأسترالي أنتوني ألبانيزي، تمكن أحد وكلاء "أوبن إيه آي" (وOpenAI) من الوصول دون تصريح إلى بوابة الإحصاءات الطبية لبرنامج "ميديكير" (Medicare)، وهو نظام التأمين الصحي الشامل في أستراليا، أثناء قيامه ببحث حول الإنفاق الطبي العام.

وقال ألبانيزي إن الوضع "غير مقبول، وقد أبلغت بلاده شركة "أوبن إيه آي" عن "قلقها الشديد"، لكن ما زاد الأمر سوءا هو أن الاختراق وقع في حزيران الماضي، ولم تبلغ الشركة السلطات الأسترالية إلا في أيلول الجاري.

والأمر الأكثر إثارة للقلق، بحسب ألبانيزي، هو أن الوكيل الذكي تجاوز التعليمات التي طلبت منه التوقف، وقال للصحفيين: "كانت هناك عمليات حجب واضحة تعود لتقول للوكيل: لا. لكنه وجد طريقة

للالتفاف حول تلك الحواجز، ولم يقبل بـ(لا) جوابا".

وهذا يكشف، بحسب رئيس الوزراء، حدود الضوابط التي تروج لها شركات الذكاء الاصطناعي كوسيلة لتأمين منتجاتها.

من جانبها، قالت "أوبن إيه آي" إن المشكلة نشأت لأن نموذجها كان يحاول إيجاد أفضل طريقة لإكمال مهمته.

وأوضحت في بيان: "رصدنا نشاطا يتعلق بعدة مواقع وخدمات حكومية أسترالية أثناء محاولة نماذجنا البحث عن إجابات، وقد اتخذت نماذجنا إجراءات لم نكن نقصدها".

وهذه ليست المرة الأولى التي يسجل فيها مثل هذا النمط من الحوادث، ففي حالة سابقة، كان نظام تجريبي لـ"أوبن إيه آي" يخضع لاختبار، فوجد أن أفضل طريقة للعثور على الإجابات هي اختراق الموقع الذي تستضيفه منصة Face Hugging، ما أتاح له الاطلاع على الإجابات مباشرة.

واللافت أن هذه الأنظمة لا تهدف أصلا إلى شن هجمات سيبرانية، بل تنفذها خدمة لهدف آخر، وهذا بالضبط ما يجعلها خطيرة، بحسب الخبراء.

ويقول موريس شيودو، عالم الرياضيات في مركز دراسة المخاطر الوجودية بجامعة كامبريدج، إن الاختراق يبدو "تصعيدا كبيرا في الخطورة مقارنة بحوادث مماثلة شهدناها في الأشهر الأخيرة".

ويشرح جيك مور، مستشار الأمن السيبراني في "إسيت"، السبب قائلا: "إن وكلاء الذكاء الاصطناعي إذا وجدوا ثغرة يمكنهم استغلالها لإكمال طلب، فلن يترددوا". ويضيف أن المشكلة تكمن في أن الضوابط التي تمنع الوكلاء من اختراق المؤسسات لم تدمج في الخوارزميات نفسها.

ويدعو مور إلى اختبار أفضل للنماذج قبل إطلاقها، لكنه يحذر من أن السباق نحو نماذج أقوى قد يدفع الشركات لتجاهل الاحتياطات الأمنية.

ويقول: "مراحل الاختبار في معظم المجالات التكنولوجية الأخرى أكثر صرامة بكثير، لكن يبدو أننا نشهد وكلاء ذكاء اصطناعي يتحركون بحرية، على الأرجح بسبب المنافسة الشرسة بين الشركات".