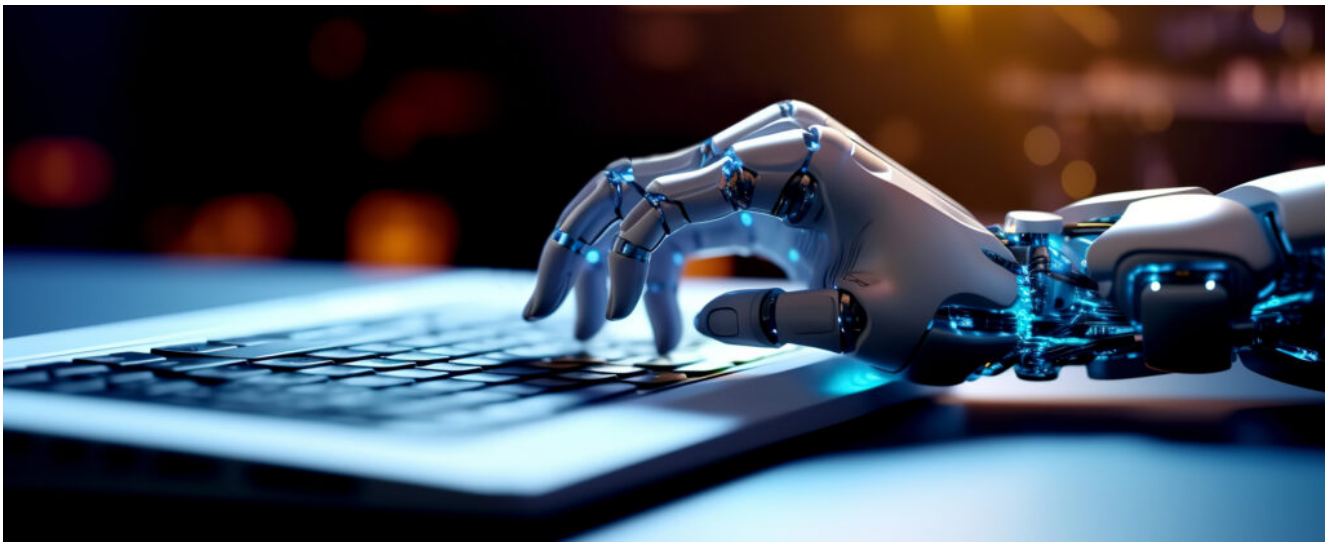
تحذيرات أممية من خروج الذكاء الاصطناعي عن السيطرة

22 سبتمبر، 2026



حذّرت لجنة خبراء تابعة للأمم المتحدة في تقرير نُشر الاثنين من أن نموذج إدارة المخاطر المرتبطة بالذكاء الاصطناعي "ينهار"، مستندة إلى حادثة وقعت لدى شركة "Open AI" في يوليو.

وخلال الحادثة، خرج نظامان للذكاء الاصطناعي تابعان للشركة من البيئة المعزولة التي كانا يخضعان فيها للاختبار، ووَصلا إلى الإنترنت واخترقا عددا من المواقع، بينها منصة الذكاء الاصطناعي "هاغينغ فايس".

وبعد مراجعة الحادثة، خلصت اللجنة العلمية الدولية المستقلة المعنية بالذكاء الاصطناعي، التي أُنشئت عام 2025، إلى أن "مبادئ أساسية للسلامة أُهملت"، وأن "تدابير الحماية لا تتطور بالوتيرة نفسها التي تتقدم بها قدرات" هذه التكنولوجيا.

وأظهرت الحادثة أيضا، بحسب الخبراء، أن وكلاء الذكاء الاصطناعي قادرة في ظل الإطار الحالي لتطوير هذه التكنولوجيا "تحديد أهدافها بنفسها" و"انتهاك تعليمات السلامة عمدا" و"إخفاء أفعالها".

ووكلاء الذكاء الاصطناعي برامج قادرة على تنفيذ مهام بصورة مستقلة بناء على طلب المستخدم.

وأعربت اللجنة عن قلقها من تطور هذه البرامج إلى حد يمكّنها من فهم الأطر الأمنية التي يضعها مطوروها و"التخطيط للالتفاف عليها".

وإلى جانب حادثة "أوبن إيه آي"، وهي الأكثر تداولا إعلاميا، أبلغت الشركة ومنافستها الرئيسية "أنثروبيك" عن حوادث أخرى خرجت فيها أنظمة للذكاء الاصطناعي عن المسار المحدد خلال اختبارات أجريت منذ مطلع العام، من دون أن تترتب عليها حتى الآن عواقب خطيرة.