

Raport: AI poate produce erori sau brese de securitate cu efecte financiare imediate, când ajunge sa execute tranzactii

Inteligența Artificială (AI) nu se mai limitează la recomandări și ajunge să execute tranzacții, iar erorile sau bresele de securitate pot produce efecte financiare imediate, se arată într-un raport de specialitate, dat publicității miercuri.

Conform analizei realizată de specialiștii Bitget și SlowMist, odată cu migrarea de la recomandare la acțiune efectuată de AI, apare și o nouă categorie de risc, pentru care modelele tradiționale de securitate nu au fost concepute încă.

"În momentul în care Inteligența Artificială nu se mai limitează la recomandări și ajunge să execute tranzacții, erorile sau bresele de securitate pot produce efecte financiare imediate. În piața crypto, unde tranzacțiile sunt procesate aproape instant, un agent compromis sau directionat greșit poate lua decizii înainte ca oamenii să poată interveni. Sistemele AI autonome deschid noi puncte vulnerabile în privința securității, de la datele pe care le primesc până la modul în care execută tranzacții. Concret, aceste sisteme pot fi influențate de comenzi înșelătoare, aplicații periculoase sau API-uri prea permissive, ceea ce poate duce la acțiuni nedorite. Astfel, pentru ca acest tip de agenți funcționează continuu, fără supraveghere din partea utilizatorului, riscul ca o eroare sau un atac să producă efecte negative este și mai mare", se menționează în raport.

În opinia experților, aceste vulnerabilități nu sunt cazuri izolate, ci fac parte dintr-un risc sistemic.

"Asadar, în această nouă etapă, securitatea nu mai poate fi gândită doar ca măsuri de protecție la nivelul aplicației, ci trebuie integrată în întreaga arhitectură prin care sistemele AI interacționează cu capitalul. Pe fondul acestor riscuri, platformele care folosesc agenți AI sunt nevoite să regândească măsurile de siguranță adoptate, subliniază raportul menționat. Una dintre soluțiile tot mai des invocate este separarea clară între analiză, execuție și accesul la fonduri, pentru a evita situațiile în care o singură vulnerabilitate poate declanșa tranzacții neintenționate. Permisunile sunt structurate pe principiul accesului minim necesar, iar înainte ca execuția să fie finalizată sunt introduse procese de simulare și verificare a tranzacțiilor. Aceste controale sunt concepute pentru a garanta ca, deși agenții AI operează autonom, aria lor de acțiune rămâne clar definită și limitată", notează sursa citată.

Potrivit raportului, este nevoie de un model de securitate care să funcționeze în toate etapele unei tranzacții, adică înainte, în timpul și după execuție.

În același timp, monitorizarea continuă, permisunile limitate și verificarea fiecărei tranzacții reprezintă elemente de bază într-un asemenea cadru, iar securitatea nu mai este doar o reacție, ci un sistem integrat.

"Datele indică o realitate mai amplă, în care agenții AI devin tot mai integrați în activitățile de tranzacționare, administrare a activelor și operațiuni on-chain, iar granița dintre intenția utilizatorului și execuția sistemului devine din ce în ce mai abstractă. În acest context, gradul de încredere nu mai este determinat exclusiv de performanță, ci și de capacitatea sistemelor de a funcționa cu limite controlate. Activitatea financiară devine tot mai automatizată, însă infrastructura trebuie să fie concepută nu doar pentru viteză și acces, ci și pentru control și reziliență", susțin specialiștii.

Raportul realizat de Bitget și SlowMist este un punct de referință pentru platformele de tranzacționare cu active digitale, dezvoltatori și utilizatori care navighează această tranziție.