USA tühistab Anthropicu ekspordipiirangud

4 kuud tagasi - 01.07.2026 Autor: AM

USA kaubandusministeerium tühistas täna Anthropicu tehisintellektimudelite ekspordipiirangud ja ettevõte alustas taas Fable 5 ning Mythos 5 mudelite pakkumist kõigile.

Anthropic teatas, et Fable 5 jõuab alates tänasest üle maailma Claude'i platvormidel kasutusse ning hiljem lisandub saadavus AWS-is, Google Cloudis ja Microsoft Foundrys. Ettevõte töötab ka USA valitsusega, et laiendada juurdepääsu Mythos 5 mudelile väljapoole praegust heakskiidetud organisatsioonide rühma Project Glasswingis.

Otsus tuli vähem kui kolm nädalat pärast seda, kui USA presidendi Donald Trumpi administratsioon käskis Anthropicul peatada kõige võimsamate mudelite kasutus välismaalastel. Toona viidati riikliku julgeoleku muredele, mis seostusid võimaliku kaitsesüsteemide läbimurdega.

Uued kaitsed

Anthropic ütles, et piirangud tühistati pärast uute kaitsemeetmete kasutuselevõttu, mis peaksid takistama kasutajaid Fable 5 küberturvalisuse kaitsetest mööda hiilimast. Ettevõte töötas raporti kontrollimiseks koos valitsusasutuste ja partneritega, sealhulgas Amazoniga, ning arendas uue klassifikaatori, mis blokeerib tuvastatud tehnikat enam kui 99% juhtudest.

Muudatus võib suurendada valepositiivseid tulemusi, kuid ettevõtte hinnangul oli see vajalik, et muud võimalused laiemalt kättesaadavaks teha. Anthropic rõhutas ka, et algsed probleemid ei paljastanud Mythos 5-le omaseid ainulaadseid kübervõimekusi.

Mis saab edasi?

Ettevõte kavatseb süvendada koostööd USA valitsusega piiriülese tehisintellekti turvalisuse teemal. Plaanide seas on varasem ligipääs uutele mudelitele, kiirem teabevahetus olulisemate jailbreak-juhtumite kohta ning ühised uurimistööd tehisintellekti ohutuse vallas.

Anthropic pakkus välja ka standardiseeritud raamistiku tehisintellekti jailbreak’ide tõsiduse hindamiseks. Ettevõtte sõnul puudub sellisel hindamisel praegu AI-tööstuses ühine arusaam ning sama raamistikku arendatakse koos Amazoni, Microsofti, Google’i ja teiste Glasswingi partneritega.

- Uudised
- Tehisintellekt

Pilt

