Euroopa Liit võtab tehisaru kontrolli alla

2 aastat tagasi - 06.02.2024 Autor: AM

Euroopa Liidu liikmesriikide saadikud jõudsid möödunud nädalal kokkuleppele tehisaru määruses, mille eesmärgiks on tagada tehisaru ohutu ja vastutustundlik kasutamine. Tegemist on maailmas esimese teedrajava õigusaktiga, mis reguleerib nii tehisaru süsteemide arendamist kui ka nende turuleviimist, pakkumist ja kasutamist, toob välja advokaadibüroo Hedman vandeadvokaat Toomas Seppel.

Tehisintellekti määruse nõuete rikkumise eest kehtestatakse ka miljonitesse ulatuvaid trahve. Näiteks võib nõuete rikkujat tulevikus oodata trahv kuni 7,5 miljonit eurot või 1,5% ettevõtte aastakäibest. Suurte globaalsete ettevõtete puhul võivad trahvisummad ulatuda koguni 35 miljoni euroni ning moodustada kuni 7% ülemaailmsest käibest. “Trahvide suurused sõltuvad nii rikkumise raskusest kui ka ettevõtte suurusest ja käibest. Selline tugev kontrollimehhanism on oluline, et ennetada rikkumisi ja tagada, et määrus kaitseks inimeste õigusi tõhusalt, edendades vastutustundlikku tehisintellekti arendamist ja kasutamist,” selgitas Seppel.

Inimeste jaoks on aga oluline määrusega kaasnev läbipaistvus - inimestega suheldes peavad tehisaru süsteemid kasutajat teavitama, et nad suhtlevad masinaga. See puudutab näiteks automaatset isikutuvastust piirikontrollis ja juturoboteid klienditeeninduses, Snapchatis ja ChatGPTs.

Lisaks peavad inimesed, kes kasutavad tehisaruga genereeritud süvavõltsinguid ehk nn deepfake, märkima, et sisu on genereeritud AI abil. Eraisikutele küll trahve ette ei nähta, kuid see ei välista deepfake´i kasutaja vastutust üldistel alustel näiteks kellegi laimamise korral.

Määrus toob esmakordselt kaasa üldotstarbeliste tehisaru süsteemide mõiste ehk GPAI (general purpose artificial intelligence). Nendest tuntumad on täna ChatGPT, Bard ja DALL-E ehk need süsteemid, mis oskavad täita paljusid erinevaid ülesandeid nagu näiteks tekstide, piltide või helide genereerimine. GPAI süsteemid peavad tulevikus vastama läbipaistvusnõuetele. Määrus reguleerib ka GPAI võimsamaid mudeleid, mis võivad põhjustada süsteemset riski. Nendest tuntumad on OpenAI (GPT-3, GPT-4), DeepMind (AlphaGo, AlphaFold) ja IBM (IBM Watson) ning nendele hakkavad kehtima lisakohustused.

Madala, kõrge ja keelatud riskiga tehisaru süsteemid

Hedmani vandeadvokaat selgitas, et määrus liigitab tehisintellekti süsteemid kolme kategooriasse vastavalt potentsiaalsetele riskidele: madala, kõrge ja keelatud riskiga.

Lubatud tehisintellekti süsteemide suhtes kehtivad üldised läbipaistvusnõuded, milleks on näiteks tehnilise dokumentatsiooni koostamine, autoriõiguse nõuete järgimine ja treenimise käigus kasutatud andmete kohta kokkuvõtete avalikustamine.

Suure mõjuga generatiivsete süsteemide ja keelatud riskiga süsteemide suhtes kehtestatakse lisakohustused ja keelud.

Kõrge riskiga tehisaru süsteemidele kehtestatakse ulatuslikud kohustused, näiteks riskijuhtimissüsteemi loomine, tehnilise dokumentatsiooni koostamine ja kaasajastamine, läbipaistvusnõuete täitmine ning inimliku järelevalve tagamine. Kõrge riskiga tehisaru süsteemid on näiteks sellised, mida kasutatakse elutähtsas taristus, meditsiiniseadmetes ning õiguskaitse ja õigusemõistmise käigus.

Euroopa Liit jõudis üksmeelele, et inimese põhiõigustele ohtu kujutavad tehisaru süsteemid liigitatakse keelatud riskiga süsteemideks ning nende arendamine ja kasutamine Euroopa Liidus keelatakse. Keelatud süsteemide seas on näiteks tundlikke andmeid kasutavad biomeetrilised kategoriseerimissüsteemid ja näotuvastuse andmebaasid. Keelatud on näiteks emotsioonituvastus töökohal ja haridusasutustes ning sellised süsteemid, mis manipuleerivad inimese alateadlikku käitumist ja kasutavad ära inimeste haavatavusi.

Eesti arendab AI määrusest lähtuvalt oma tehisaru strateegia

Tehisaru määrus ootab nüüd edasist täpsustamist ja vastuvõtmist ELi seadusandjate poolt. Ka Eesti arendab hetkel AI määrusest lähtuvalt oma tehisaru strateegia, mis annab juhtnöörid ja loob standardid ning vaatab süsteemsemalt, kuidas tagada AI usaldusväärsus ning maandada riske nii AI arendades kui ka kasutades.

Tehisaru määrus kiidetakse eeldatavasti komiteede tasandil heaks järgmise kahe nädala jooksul. Seejärel suunatakse see ELi täiskogu hääletusele Euroopa Parlamendis, mis peaks toimuma aprillis. Täpsemalt saab määrusest lugeda advokaadibüroo Hedman kodulehelt.

- Tegijad

- Uudised

- Lahendused
- Tarkvara

Pilt

