OUCH! 2026. gada jūlijs

Workforce Security and Risk Training

Padomājiet, pirms izmantojat mākslīgo intelektu (MI): kā to darīt droši

Vienreizēja iespēja, kas tika nozagta

Lēna nesen sāka izmantot mākslīgā intelekta sarunu botu, lai vieglāk tiktu galā ar savu aizņemto ikdienu. Viņai patika, cik viegli ir uzdot jautājumus un uzreiz saņemt atbildes par ģimeni, rēķiniem un nākotnes plānošanu. Kādu vakaru, jūtot satraukumu par savām finansēm, viņa lūdza MI palīdzību investīciju jautājumos. Tas ātri ieteica stratēģiju. Tomēr mākslīgajam intelektam nebija pieejama pilnīga informācija par Lēnas finansiālo situāciju, riska toleranci un nodokļu ietekmi tieši viņas konkrētajā gadījumā. Sarunu bots ieteica pārvietot naudu uz populāru akciju un īstermiņa darījumu kombināciju, kas solīja lielāku peļņu. Tāpat paskaidroja, kā rīkoties ar nodokļiem saistībā ar viņas investīcijām.

Padoms izklausījās pārliecinošs un pārdomāts, tāpēc Lēna tam uzticējās. Sākumā viņa bija sajūsmā. Taču dažos mēnešos tirgus mainījās un viņas investīcijas zaudēja vērtību. Situācija pasliktinājās, kad pienāca nodokļu deklarāciju iesniegšanas laiks, viņa atklāja, ka ir pārpratusi svarīgus noteikumus. Tā kā viņa bija rīkojusies saskaņā ar mākslīgā intelekta norādījumiem, tos nepārbaudot, viņa pieļāva kļūdas nodokļu aprēķinos, kas noveda pie tā, ka bija jāmaksā soda naudas un radās arī vēl citas papildu izmaksas.

Rezultātā uzticoties padomam, kas viņas situācijā nebija piemērots, Lēna zaudēja naudu. MI var būt noderīgs rīks, taču ir svarīgi ņemt vērā, ka tas var arī kļūdīties. Ja paļaujaties uz to bez papildus pārbaudes, nelielas kļūdas var drīz vien pārvērsties problēmās, kas var dārgi maksāt.

Kas ir mākslīgais intelekts (MI)?

Mākslīgais intelekts (MI) ir tehnoloģija, ir kas izstrādāta, lai simulētu cilvēku domāšanu, informācijas apstrādi un lēmumu pieņemšanu. Tas var iekļaut valodas ģenerēšanu, attēlu atpazīšanu, lēmumu pieņemšanu, satura radīšanu vai problēmu risināšanu. Kopumā ir trīs MI veidi, kurus ir iespējams izmantot.

- **Integrētais MI**: tas ir MI, kas ir iebūvēts rīkos, kurus izmanto ikdienā, bieži vien MI tiek izmantots, pašam to neapzinoties. Piemēram, kad fotografējat ar tālruni, MI, visticamāk, uzlabo attēlu.

- **Ģeneratīvais MI**: tie ir specializēti MI pakalpojumi, kas ir paredzēti cilvēku atbalstam, tostarp tādi rīki kā ChatGPT, Google Gemini vai Anthropic Claude. Ģeneratīvais MI var palīdzēt daudzos apstrādes uzdevumos, piemēram, mūzikas radīšanā, biznesa plāna rakstīšanā, attēlu ģenerēšanā vai ideju analizēšanā.

- **Aģentos balstīts (Agentic AI) MI**: šie ir mākslīgā intelekta pakalpojumi, kas paredzēti, lai veiktu darbības jūsu vārdā. Šīs sistēmas var darboties ar ierobežotu cilvēka iesaisti vai pat bez tās un funkcionēt kā daļa no digitālā darbaspēka, pieņemot lēmumus un veicot darbības, pamatojoties uz vispārīgām vadlīnijām vai norādījumiem.

Mākslīgā intelekta droša lietošana

No visiem trim MI veidiem pievērsīsimies ģeneratīvajam MI (GenAI), jo tas ir MI veids, kuru, visticamāk, izmantojat visbiežāk. GenAI ir jaudīgs rīks, kas var palīdzēt efektīvāk paveikt uzdevumus un apgūt jaunas prasmes un idejas, ja to izmantojat droši. Lūk, daži aspekti, kas ir jāņem vērā.

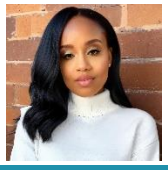
- **Privātums**: esiet piesardzīgi attiecībā uz to, kādu informāciju sniedzat MI rīkiem. Augšupielādējot vai ievadot jebkādu informāciju MI sistēmā, tā var tikt apstrādāta un saglabāta, un dažos gadījumos – izmantota pakalpojuma uzlabošanai atkarībā no platformas. Ja ar MI rīkiem dalāties īpaši sensitīvu informāciju, tā var kļūt pieejama citiem. Pirms dalīšanās ar informāciju, apdomājiet to, vai jūs justos droši, ja to zinātu visa pasaule. Vēl viena iespēja ir izmantot maksas MI pakalpojumus, kas aizsargā privātumu, izmantojot modeļus, kas nemācās no jūsu datiem.

- **Precizitāte**: MI var apgalvot, ka tā radītais vai attēlotais saturs ir precīzs, pat ja tas ir kļūdains. Vienmēr rūpīgi pārbaudiet un pārliecinieties par mākslīgā intelekta sniegto rezultātu. Izplatīta MI problēma ir tā, ka tas vienmēr centīsies sniegt atbildi, pat ja nesaprot jautājumu. Pēc noklusējuma MI nelūgs precizēt savus pieprasījumus, tāpēc ir svarīgi būt pēc iespējas konkrētākam un pievērst uzmanību neskaidriem vai mulsinošiem rezultātiem MI atbildēs.

- **Priekšstati**: tāpat kā cilvēkiem, kuri ir programmējuši MI, arī MI var būt savi priekšstati Tas var sniegt atbildes, kas izklausās pārliecinošas, bet nav izsvērtas vai precīzas. Ļoti izplatīta tendence ir tā, ka MI vēlas "iepriecināt" lietotāju, tāpēc tas mēdz teikt to, ko, pēc tā domām, vēlaties dzirdēt. MI “zina” tikai to informāciju, ar kuru tas ir mācīts un kurai tam ir piekļuve; ja šī informācija ir būtiski kļūdaina vai ierobežota, arī atbildes atspoguļos šos trūkumus.

MI ir viens no spēcīgākajiem rīkiem, kas šobrīd ir pieejams. Tas var palīdzēt strādāt ātrāk, uzzināt vairāk un būt produktīvākam. Taču, tāpat kā jebkurš spēcīgs rīks, tas ir jāizmanto ar apdomu Tam nevajag akli ticēt. Apdomājiet, cik daudz un kāda veida informāciju tam sniedzat. Nepiešķiriet tam lielāku kontroli, nekā nepieciešams. Izmantojiet MI kā rīku, kas palīdz pieņemt lēmumus, nevis aizstāj jūsu spriedumu.

Viesredaktore



Portija Džefersone (Portia Jefferson) ir kiberdrošības speciāliste, kura koncentrējas uz MI drošību, risku apzināšanu un praktiskām vadlīnijām ikdienas lietotājiem. Ar pieredzi finanšu tehnoloģiju un nodokļu jomā viņa palīdz cilvēkiem droši orientēties jaunajās tehnoloģijās darbā un mājās.

Resursi

Uzmanieties no dziļviltojumiem. Jauns maldināšanas laikmets: https://www.sans.org/newsletters/ouch/beware-deepfakes-new-age-of-deception
Iedomātās balsis: aizsardzība pret balss klonēšanas uzbrukumiem: https://www.sans.org/newsletters/ouch/phantom-voices-defend-against-voice-cloning-attacks
Sevis pasargāšana, ja īsta privātuma sasniegšana nav iespējama: https://www.sans.org/newsletters/ouch/protecting-yourself-when-true-privacy-is-impossible

Tulkoja: CERT.LV

Vairāk informācijas par Ouch! Varat atrast šajā saitē: https://www.sans.org/newsletters/ouch