

OUCH!

Mėnesinis informacinio saugumo naujienlaiškis Tau

Dirbtinis intelektas: Ką verta žinoti?

Kas tai yra ir kodėl turėtų man rūpėti?

Dirbtinis intelektas (DI) apibūdina sistemas, užprogramuotas mąstyti ir reaguoti kaip žmonės. Tiesą sakant, mes uždavėme DI sprendimui ChatGPT (pokalbių robotas) tą patį klausimą ir gavome šį atsakymą.

Kas yra dirbtinis intelektas?

Dirbtinis intelektas yra bendras terminas, apibūdinantis įvairius metodus, leidžiančius mašinoms ar kodui imituoti žmogaus intelektą. Dirbtinis intelektas leidžia techninėms sistemoms suvokti savo aplinką, susitvarkyti su tuo, ką suvokia ir išspręsti problemas siekiant konkretaus tikslo. Yra keletas dirbtinio intelekto tipų, įskaitant pagrįstas taisyklės, ekspertines sistemas ir mašininį mokymąsi.

Dirbtinis intelektas įgauna tokį galingumą, kad jis gali imituoti žmogaus proto intelektą ir mąstymo gebėjimus, tačiau jis gali išanalizuoti eksponentiškai daug daugiau informacijos nei bet kuris žmogus ir tai padaryti eksponentiškai greičiau.

Dirbtinio intelekto samprata nėra nauja. Istoriniuose šaltiniuose galima rasti įvairių mitų ir pasakojimų apie geležinius žmones, su dirbtinėmis smegenimis, kurie buvo protingesni už bet kurį kitą gyvą žmogų. Tačiau tokie pasakojimai liks pasakojimais, nes nėra faktų, paremtų įrodymais. Priežastis, dėl kurios dabar tiek daug apie tai girdime, yra ta, kad pirmą kartą kiekvienas turi galimybę bendrauti ir pamatyti tikrąjį dirbtinio intelekto funkcionalumą.

ChatGPT (pokalbių robotas), internetinis dirbtinio intelekto pokalbių robotas, yra vienas iš pirmųjų viešai prieinamų sprendimų, galinčių reaguoti kaip tikras žmogus, išlaikęs tai, kas vadinama Turingo testu. Šis testas nustato mašinos gebėjimą parodyti protingą elgesį, kai tikras žmogus sąveikauja su mašina per tekstinį pokalbių srautą. Jei žmogus negalėjo pasakyti, ar jis sąveikauja su mašina ar asmeniu, mašina išlaikė testą. Dirbtinio intelekto sprendimai šiandien yra pirmieji viešai prieinami.

Tačiau internetiniai pokalbiai yra tik pradžia to, ką dirbtinis intelektas gali atlikti. Dabar yra dirbtinio intelekto sprendimų, kurie gali sukurti vaizdo įrašą, kuriame asmuo veda pamoką, bet kuria kalba, analizuoti sveikatos įrašus ir greitai nustatyti, kas greičiausiai suserga vėžiu, kurti naujienų straipsnius pagal pasirinktą temą, kurti vaizdus vaikiškoms knygelėms, arba sukurti naują kompiuterių programų kodą. Nors dirbtinis intelektas nebūtinai yra kažkas tokio, ko reikia bijoti, yra keletas pavojų, kuriuos reikia žinoti.

Dirbtinio intelekto pavojai

1. **Atkuriant asmenį:** Dirbtinio intelekto sprendimai gali sukurti Jūsų balso ir tono kopiją ir ją naudoti realiuoju laiku sukurdamas garsą, kuris bus atkartojamas kaip Jūsų balsas. Tai kelia riziką ir abejones šia technologija. Taigi, kibernetinis užpuolikas gali įrašyti telefono balso pranešimą, kuris skambės Jūsų balsu, taip siekdamas apgauti Jūsų bendradarbius, banką ar šeimos narį su tikslu imtis tam tikros nusikalstamos veikos. Dirbtinis intelektas ta patį gali sukurti pasinaudodamas nuotraukomis ir vaizdo medžiaga. Deep Fakes (Mašininis mokymasis) dirbtinio intelekto tipas naudojamas kuriant įtikinamus vaizdus, garso ir vaizdo apgaules. Taip sukuria netikrą turinį.
2. **Klaidingi atsakymai:** Kalbant apie dirbtinio intelekto teikiamus duomenis ar atsakymus, sprendimai gali būti klaidingi. Dirbtinis intelektas dažnai naudoja viešą informaciją, bet atsakymams gali turėti įtakos kūrėjų šališkumas. Nors įprasti paieškos varikliai yra sukurti taip, kad pateiktų Jums geriausią arba teisingiausią atsakymą į Jūsų užklausas, tokie sprendimai kaip dirbtinio intelekto gali būti sukurti taip, kad pateiktų labiausiai žmogišką atsakymą. Kuris yra geresnis, priklauso nuo to, ką bandote pasiekti.
3. **Ne visi lygūs!** Dirbtiniam intelektui tapus naujausia technologija, dabar yra šimtai pradedančiųjų įmonių, siūlančių įvairias dirbtinio intelekto paslaugas. Daugelis iš jų nori Jūsų informacijos ar kredito kortelės bandomajam naudojimui. Būkite atsargūs – ne visos dirbtinio intelekto paslaugos yra patikimos. Atlikite savo tyrimus prieš prisiregistruodami ir naudodamiesi dirbtinio intelekto paslauga.
4. **Privatumas:** Kai naudojate dirbtinio intelekto sistemą arba bendraujate su ja, pvz. bendraudami su ChatGPT, atminti, kad bet kokią informaciją, kurią įvesite į sistemą, bus ne tik apdorojama, bet ir išsaugoma bei panaudojama atsakymams kitiems asmenims. Tai reiškia, kad jeigu įvesite bet kokią asmeninę informaciją apie save arba bet kokią kitą konfidencialią informaciją, tai ta informacija bus saugoma ir gali būti bendrinama arba parduota kitiems. Darbe nesidalykite ir neveskite jokios informacijos, kuri, Jūsų manymu, yra neskelbtina, asmeninė ar konfidenciali.

Dirbtinio intelekto ateitis

Dirbtinis intelektas vis dar yra labai ankstyvoje stadijoje, panašiai kaip internetas buvo prieš dvidešimt ar trisdešimt metų. Nors galime tikėtis greitos dirbtinio intelekto evoliucijos ir pritaikymo, labai sunku numatyti, koks bus jo poveikis. Tiesiog atminti, kad šios galimybės egzistuoja ir naudodamiesi dirbtiniu intelektu būkite labai atidūs, kokią informaciją įvedate ir bendrinate.

Šaltiniai

ChatGPT: <https://chat.openai.com/chat>

Turingo testas: https://en.wikipedia.org/wiki/Turing_test

Išvertė UAB Perlas Network

OUCH! Skelbiamas SANS saugumo supratimas ir platinamas pagal Kūrybingi panašumai BY-NC-ND 4.0 licenciją. Šį naujienlaiškį galima laisvai dalytis ir platinti tol, kol nebus pakeistos sutarties sąlygos. Redakcinė kolegija: Walter Scrivens, Phil Hoffman, Alan Waggoner, Leslie Ridout, Princess Young.