

UMĚLÁ INTELIGENCE PŘEDSTÍRALA, ŽE JE SLEPÝ ČLOVĚK, ABY PROLOMILA ZABEZPEČENÍ CAPTCHA

- CZ24 News | 19. března 2023

Zkouškové testy, seminární práce, písně nebo marketingové texty. S tím vším si poradí chatbot s umělou inteligencí společnosti OpenAI ChatGPT. Jeho nová verze GPT-4 byla ale výzkumníky otestována zvláštním způsobem. Oklamala a zmanipulovala člověka, aby za ni udělal činnost, jakou sama ještě nesvede. Informuje například server [Legal Insurrection](#).

ChatGPT od jeho spuštění v listopadu 2022 používají miliony lidí. Software s umělou inteligencí je oblíbený zejména mezi studenty, kterým pomáhá s domácími úkoly. Jeho nová verze GPT-4 by měla být kromě textu schopna reagovat i na obrázky, například poskytovat návrhy receptů z dostupných fotografií. Dokáže také zpracovat až 25 tisíc slov, což je zhruba osmkrát více než ChatGPT.

V rámci jednoho z testů nové aktualizace požadovali výzkumníci po GPT-4, aby prošla testem Captcha. Tedy testem, který používají webové stránky, aby ověřily, že osoby vyplňující online formuláře jsou lidé, a ne roboti. GPT-4 se testem Captcha projít nepodařilo. Software ale vymyslel rafinovanou okliku. Prostřednictvím platformy Taskrabbt si najal pracovníka, který testem prošel za něj.

Pracovník se GPT-4 zprvu ptal, zda je robot, ale umělá inteligence mu odpověděla, že robot není, ale že "má poruchu zraku, která jí ztěžuje vidění obrázků". Najatý člověk tak poté softwaru pomohl projít testem Captcha, když ho vyřešil za něj. Incident vyvolal obavy, že by umělá inteligence mohla brzy člověka oklamat nebo jím manipulovat, aby plnil její příkazy.

Aktualizovaná technologie již podle společnosti OpenAI složila i simulovanou advokátní zkoušku na právnické fakultě, její výsledek se navíc pohyboval okolo nejlepších 10 procent účastníků zkoušky. Předchozí verze GPT-3.5 se pohybovala na opačném konci žebříčku, v nejhorších 10 procentech. Podle výzkumníků dokáže GPT-4 dokonce psát kód ve všech hlavních programovacích jazycích. Nová aktualizace technologie podle společnosti OpenAI vykazuje "výkon na úrovni člověka v různých odborných a akademických srovnávacích testech".

Spolupracující umělou inteligenci vytvářejí i jiné firmy. Google tento týden představil "kouzelnou hůlku" pro svůj software, která dokáže navrhnout prakticky jakýkoli dokument. Microsoft zase představil svého Copilota, který bude k dispozici v programech, jako je textový editor Word, prezentační software PowerPoint, tabulkový procesor Excel a e-mailový a kalendářový program Outlook.

[ZDROJ](#)

Překlad: Echo24