

„POKUD BUDEME TAKTO POKRAČOVAT, VŠICHNI ZEMŘOU,“ varuje odborník na AI a vyzývá k absolutnímu vypnutí

- CZ24 News | 6. dubna 2023

Stále více odborníků varuje před dalším pokračováním vývoje AI, neboť lidstvo by mohlo dopadnout podobně jako v některých sci-fi filmech, kde AI převzala kontrolu nad člověkem. Jde o zcela reálnou hrozbu a to potvrzuje i další odborník.

„Lidé nejsou připraveni na výkonnou AI za současných podmínek a dokonce ani v dohledné budoucnosti,“ vysvětlil přední odborník v oboru a dodal, že nedávný otevřený dopis vyzývající k šestiměsíčnímu moratoriu na vývoj pokročilé umělé inteligence ještě podcenil vážnost situace.

„Hlavním problémem není inteligence, která může konkurovat lidem (jak uvádí otevřený dopis), ale co se stane, když je umělá inteligence chytřejší než lidé,“ řekl Eliezer Yudkowsky, teoretik rozhodování a přední výzkumník umělé inteligence.

„Mnoho výzkumníků studujících tyto otázky, včetně mě, předpokládá, že nejpravděpodobnějším výsledkem vývoje nadlidsky inteligentní umělé inteligence za současných okolností je, že doslova všichni na Zemi zemřou.“

„Ne ve smyslu, že je zde možná vzdálená šance, ale ve smyslu, že je to jasná věc, která by se prostě stala. Není to tak, že byste v zásadě nemohli přežít tím, že vytvoříte něco mnohem chytřejšího, než jste vy; jen by to vyžadovalo přesnost a přípravu a nové vědecké poznatky a pravděpodobně nemáte systémy umělé inteligence složené z obrovských nevyzpytatelných polí zlomkových čísel.“

Po nedávné popularitě a explozivním růstu ChatGPT podepsalo několik vedoucích firem a výzkumníků, včetně Elona Muska a Steva Wozniaka, dopis vyzývající všechny laboratoře umělé inteligence, aby přestaly trénovat systémy umělé inteligence, které jsou výkonnější než GPT-4, aby byly okamžitě pozastaveny alespoň na dobu šest měsíců. GPT-4, vydaný v březnu, je nejnovější verzi chatbota OpenAI, ChatGPT.

Yudkowsky předpovídá, že bez pečlivé přípravy bude AI ve svých tvrzeních značně odlišná od lidí, a že jakmile si uvědomí sama sebe, nebude se starat ani o nás, ani o žádný jiný vnímající život. „Tento druh péče je něco, co byste v zásadě mohli vložit do AI, ale zatím tam nejsme a v tuto chvíli nevíme jak to udělat.“

Proto volá po absolutním vypnutí.

Bez lidského přístupu k životu bude AI jednoduše vnímat všechny vnímající bytosti jako atomy, které může použít pro něco jiného. A lidstvo může udělat jen málo, aby tomu zabránilo. Yudkowsky přirovnal scénář k 10letému dítěti, které se snaží hrát šachy proti Stockfish 15.

Žádný lidský šachista nikdy nedokázal porazit Stockfish, což je považováno za nemožný výkon.

Veterán v oboru vyzval čtenáře, aby si představili, že technologie AI se nenachází v mezích internetu.

„Představte si celou mimozemskou civilizaci, myslící rychlostí milionkrát vyšší než člověk a zpočátku omezenou na počítače – ve světě tvorů, které považují za velmi hloupé a velmi pomalé.“

Umělá inteligence rozšíří svůj vliv za periferii fyzických sítí a mohla by vytvářet umělé formy života v laboratořích, které používají řetězce DNA k výrobě proteinů.

„Konečným výsledkem vybudování všemocné umělé inteligence za současných podmínek by byla smrt každého jednotlivého člena lidského druhu a veškerého biologického života na Zemi,“ varoval.

Yudkowsky obvinil OpenAI a DeepMind – dvě z předních světových laboratoří pro výzkum AI – z toho, že na toto téma nemají žádnou přípravu ani požadované protokoly. OpenAI dokonce plánuje sladit samotnou AI s lidskými hodnotami.

„Budou spolupracovat s lidmi, aby zajistili, že jejich vlastní nástupci budou více naladěni na lidi,“ řekli OpenAI.

„Tato praxe je dostatečná k tomu, aby každý zdravý člověk zpanikařil,“ řekl Yudkowsky.

Dodal, že lidé nemohou plně monitorovat nebo detekovat systémy umělé inteligence. Vědomé digitální mozky požadující „lidská práva“ by mohly zajít tak daleko, že lidé již nebudou vlastnit systém.

„Pokud si nemůžete být jisti, zda vytváříte sebevědomou AI, je to alarmující, nejen kvůli morálním důsledkům „sebevědomé“ části, ale také proto, že nejistota znamená, že nemáte ponětí, co děláte. a to je nebezpečné a mělo by to být zastaveno.“

Na rozdíl od jiných vědeckých experimentů a postupného rozvoje znalostí a dovedností si to lidé s nadlidskou inteligencí nemohou dovolit, protože pokud se to nepodaří na první pokus, není druhá šance protože budete mrtví.

Yudkowsky řekl, že mnoho výzkumníků si je vědomo toho, že „směřujeme ke katastrofě“, ale neříkají to nahlas.

Tento postoj kontrastuje s postoji zastánců, jako je Bill Gates, který nedávno chválil vývoj umělé inteligence.

Gates tvrdil, že vývoj AI byl „stejně zásadní jako vynález mikroprocesoru, osobního počítače, internetu a mobilního telefonu. Změní způsob, jakým lidé pracují, učí se, cestují, vyhledávají lékařskou péči a komunikují spolu. Přízpůsobí se tomu celá průmyslová odvětví. Společnosti budou vynikat v tom, jak dobře je používají.“

Gates řekl, že AI může pomoci s několika progresivními problémy, včetně změny klimatu a ekonomické nespravedlnosti.

Mezitím Yudkowsky nařizuje všem subjektům, včetně vlád a armád po celém světě, aby na neurčito ukončily velké tréninky AI a uzavřely všechny velké počítačové farmy, kde se AI vyvíjí. Dodává, že umělá inteligence by se měla omezovat pouze na řešení problémů v biologii a biotechnologii a neměla by být školená tak, aby četla „texty z internetu“ nebo „dostala se na úroveň, na které začne mluvit nebo plánovat“.

Pokud jde o umělou inteligenci, žádné závody ve zbrojení se nekonají. „To, že všichni žijeme nebo umíráme společně, není politika, ale přírodní fakt.“

Yudkowsky uzavírá slovy:

„Nejsme připraveni. Nejsme na cestě k tomu, abychom byli v dohledné době mnohem vyspělejší. Pokud budeme takto pokračovat, zemřou všichni, včetně dětí, které si to nevybraly a nic špatného neudělaly. Konec.“

[ZDROJ](#)

Zpracoval: Slovanka/Necenzurovaná Pravda