

B. Bartee, Global Research: JE UMELÁ INTELIGENCIA (AI) HROZBOU PRE CIVILIZÁCIU A ĽUDSTVO?

- CZ24 News | 17. dubna 2023

Ako veľmi by sme sa mali ako civilizácia obávať umelej inteligencie za predpokladu, že chceme pokračovať v existencii? Nedávno som sedel na podcaste s Nicolasom Creedom a redaktorom The Daily Bell Joe Jarvisom, aby sme diskutovali o existenciálnej hrozbe alebo jej nedostatku, ktorú predstavuje nekontrolovaná umelá inteligencia (AI).

Joe, ktorý hrá diablovho advokáta, bol nadšený z AI ako čistého pozitíva pre ľudstvo. Nicolas a ja sme boli menej optimistickí. Všetci sme sa zhodli, že určujúcim faktorom bude spôsob, akým sa vyvíja - kým, na aké účely a s akými opatreniami, ak nejaké existujú.

K nám **v tíme Skeptic** sa teraz pripája množstvo skúsených profesionálov v oblasti AI. Jedna taká postava napríklad nedávno doslova vyzvala na bombardovanie dátových centier AI, ktoré poskytujú vstupy pre „poznávanie AI“.

Cez [futurizmus](#) :

Jeden z najhlasnejších kritikov umelej inteligencie na svete dôrazne vyzval, aby sme umelú inteligenciu nielen zastavili, ale aj militantne s ňou skoncovali - skôr, ako to ona skončí s nami...

Výskumník v oblasti strojového učenia **Eliezer Yudkowsky**, ktorý už viac ako dve desaťročia varuje pred dystopickou budúcnosťou, ktorá príde, keď dosiahneme umelú všeobecnú inteligenciu (AGI), opäť zvoní na poplach.“

Pre referenciu uvedme, že „**umelá všeobecná inteligencia**“ alebo **AGI** [je všeobecne definovaná ako](#) „reprezentácia zovšeobecnených ľudských kognitívnych schopností v softvéri tak, aby systém AGI, konfrontovaný s neznámou úlohou, mohol nájsť riešenie. Zámer systému AGI je vykonávať akúkoľvek úlohu, ktorej je ľudská bytosť schopná.“

[„Yudkowsky povedal, že hoci chváli signatárov nedávneho otvoreného listu](#) Inštitútu budúcnosti života - medzi ktoré patrí generálny riaditeľ SpaceX Elon Musk, spoluzakladateľ spoločnosti Apple Steve Wozniak a bývalý kandidát na prezidenta Andrew Yang - vyzývajúci na šesťmesačnú prestávku v rozvoji AI, aby som to bilancoval, sám to nepodpísal, lebo to nezachádza dosť ďaleko.“

[Varovný list podpísaný Elonom Muskom a ďalšími významnými verejnými osobnosťami, o ktorých sa Yudkowsky zmieňuje](#) , čiastočne znie:

Mali by sme vyvinúť neľudské mysle, ktoré by nás mohli časom prevýšiť, prekabátiť, zastarať a nahradiť nás? Mali by sme riskovať stratu kontroly nad našou civilizáciou? Takéto rozhodnutia sa nesmú delegovať na nevolených technologických lídrov.

„Výkonné systémy umelej inteligencie by sa mali vyvíjať len vtedy, keď sme si istí, že ich účinky budú pozitívne a ich riziká budú zvládnuteľné. “

Iné ťažké váhy AI zopakovali podobné pocity, vrátane „[krstného otca umelej inteligencie](#)“, **Geoffreyho Hinton**a, ktorý uviedol „malé riziko“, že by umelá inteligencia bola zmarom ľudstva.

Vrátiac sa k Yudkowského výzve do doslovných zbraní, aby zastavil vzostup AI, nastoľuje základný problém, ktorý som už naznačil inde, vytvorenia **inteligencie, ktorá prekračuje kognitívne limity ľudstva**.

Bez účinných zábran, ktoré by zabránili tomu, aby sa to stalo nedbanlivosťou voči ľudskému blahobytu alebo priamo nepriateľským voči ľudskému životu, sme vo vážnej nevýhode:

„Nie je to tak, že by ste v princípe nemohli prežiť vytváranie niečoho oveľa inteligentnejšieho, ako ste vy,“ uvažoval, „ide o to, že by to vyžadovalo presnosť a prípravu a nové vedecké poznatky a pravdepodobne nebyť systémov umelej inteligencie zložených z obrovských nevyspytateľných polí, len zlomkových čísiel“.

Podobne ako biomedicínski výskumníci, ktorí využívajú výskum ziskovosti funkcií na potlačenie vírusov, **aj inžinieri AI v súčasnosti pracujú na vytváraní stále inteligentnejšej AI nevedia, čo robia**.

Zasahujú, bezohľadne a zbytočne, do síl, ktorým nerozumejú, kedy by si obozretný postup vyžadoval predchádzajúce štúdium.

Ako sa nedávno uvádzalo na inom mieste, [AI veľmi nedávno vyvinula to, čo filozofi a biológovia nazývajú „teória mysle“](#). To znamená, že má novoobjavenú živú schopnosť zarámovať sa do stavu mysle inej osoby alebo veci a potom podľa toho strategicky konať.

Možno nie je rozumné podpísať výzvy na kinetické bombardovanie informačných skladov, ale tento vývoj by určite mal slúžiť ako dôvod na prestávku, aby sme sa popasovali so širokými dôsledkami tejto technológie.

AUTOR: Ben Bartee

[ZDROJ](#)

Překlad: S.S.