

Jak bezpečně využívat umělou inteligenci?

MIROSLAV NEČAS

Téma umělé inteligence se díky GPTChat a dalším technologiím v oblasti generativní AI stále drží v popředí zájmu. Na rozdíl od ostatních aplikací se generativní AI ukazuje jako cenná pomůcka, která je cenově dostupná a snadno použitelná.

Počet klientů těchto služeb stoupá tak strmě, až to vyděsilo i ty, již donedávna AI prosazovali. „Pozastavme dočasně vývoj AI, společnost není připravena,“ žádá Elon Musk a další. Je nasazení AI opravdu nebezpečné, nebo lze AI nasadit zodpovědným a bezpečným způsobem?

Tato otázka leží na stole již delší dobu. NATO definovalo šest principů zodpovědného využívání AI ve své strategii již v roce 2021 a je lídrem v této oblasti. Evropská komise přichází s regulací AI Act, která staví na obdobných zásadách, principy zodpovědného využívání AI definují na národní úrovni Spojené státy a Velká Británie. Jak ale ověřit, že je AI využívána v souladu s těmito zásadami? Touto otázkou se zabývám poslední rok jako vedoucí mezinárodního týmu, jehož cílem je stanovit metodiku certifikace umělé inteligence v souladu s principy zodpovědného užití definovanými NATO. Jsou to: zákonnost; (z)odpovědnost, transparentnost, spolehlivost, ovladatelnost a zamezení nežádoucímu zkreslení. Metodiku, na které pracujeme, testujeme na šesti příkladech užití, jež jsou v různém stupni zralosti implementace AI. Udělali jsme analýzu několika desítek mezinárodních standardů, které jsou relevantní pro oblast implementace AI, a na základě gap analýzy a zobecnění poznatků z konkrétních příkladů užití formulujeme doporučení pro NATO DARB (Data and AI Review Board), který začíná pracovat v tomto roce.

Hlavním závěrem studie je, že AI lze implementovat a využívat zodpovědně, tak aby byly respektovány výše uvedené principy. Jsem přesvědčený, že stejně jako se lidstvo vyrovnalo s technologickými posuny v minulosti, vyrovná se



s nástupem umělé inteligence. Byť jistě bude docházet k problémům, kterým se nevyhne nástup žádné přelomové technologie – páry, elektřiny či atomové energie. Umělá inteligence bude mít dopad do řady dalších oblastí. Například z pohledu kybernetické a informační bezpečnosti je využití AI poměrně zásadní výzvou. Ukazuje se, že ti, kteří již delší dobu poukazují na nutnost celostního přístupu k bezpečnosti, měli pravdu. S nástupem AI hranice mezi fyzickou a kybernetickou bezpečností přestává existovat. Například když auto řízené AI nezabrzdí a dojde k nehodě, je potřeba začít se ptát: fungoval hardware, na kterém běžela AI, bez vady? Dostala AI správná data ze senzorů? Vyhodnotila AI situaci správně? Vydala AI pokyn k zabr-

zdění včas? Obdržel řídicí systém vozu pokyn zabrzdit? Fungovaly brzdy správně? Byly pneumatiky v pořádku? A co silnice a povětrnostní podmínky? Je patrné, že šest výše uvedených principů je navzájem propojeno: například bez transparentnosti a ovladatelnosti AI není možné vyvodit konkrétní odpovědnost.

Ukazuje se, že klíčem k zodpovědnému používání AI je pochopení role AI v konkrétním případě užití, identifikování a řízení z toho plynoucích rizik. Co je důležité, rizika spojená s AI nelze řešit odděleně od rizik spojených s kybernetickou či fyzickou bezpečností. Selhání informační technologie může být problémem, stejně jako využívání cloudové AI služby může ohrozit důvěrnost informací či vyvolat kybernetický incident. Využití AI v kybernetické bezpečnosti ať již jako bezpečnostního prvku, nebo jako vektoru útoku (například v rámci sociálního inženýrství) je kapitolou samou pro sebe.

Technologie Tovek je díky své všestrannosti ideálním pomocníkem při analýze rizik spojených s AI i při vyšetřování bezpečnostních incidentů spojených s umělou inteligencí. Dokáže propojit a vyhodnotit data z informačních technologií s daty z fyzického světa i umělé inteligence. Díky unikátní zkušenosti v oblasti certifikace umělé inteligence ve spojení s třiceti lety praxe v oblasti analýzy nestrukturovaných dat nabízíme unikátní řešení, které vám umožní aplikovat umělou inteligenci tak, abyste byli konkurenceschopní, a zároveň tak učinit zodpovědným způsobem, který vám zajistí dlouhodobou bezpečnost.

Nástroje Tovek vám pomohou vyhodnotit, zda umělá inteligence pracuje správně.

Autor je Business Development Manager společnosti TOVEK.