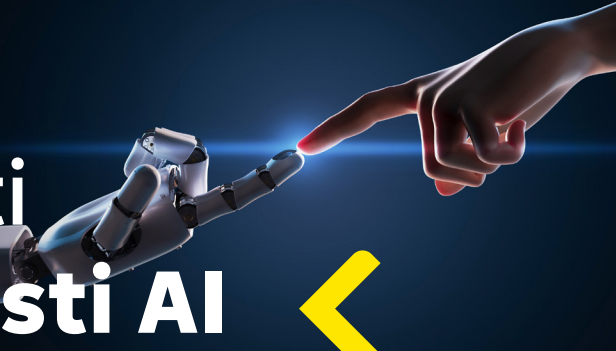




Slasti a pasti AI

Zneužívání jazykových modelů při konstrukci nápaditých podvodů v oblasti sociálního inženýrství je nálehavé téma, ale v jeho stínu se rysuje téma ještě nálehavější – **útoky na mechanismy a modely** jako takové. A právě těm se budeme věnovat. Řeč bude o technikách prompt injection, o hrozbách vykradení uživatelských dat, unesení modelu samotného nebo jeho „otrávení“. Zapomeňme na populární zkratku AI = ChatGPT.

Kdy?

27. 2. 2024 od 17:30

Kde?

Kavárna Šestnáctka

Budova City Tower,
Hvězdova 1716/2b
Praha 4 – Pankrác



TOMÁŠ ROSA
Mathematical
Security Architect

Tomáš Rosa, náš matematický expert na bezpečnost, vás zavede do reality, jak mohou tyto sofistikované útoky zvenčí ohrozit AI pracující ve službách banky, například v automatizované detekci bankovních podvodů. „Uvažujme, jako by to byl člověk, jednejme s tím jako se strojem.“ Tuhle poučku si už teď запиšte za uši. Tam vám ji nikdo nehackne.

- *Tento text byl napsán člověkem. Ale čemu dneska můžete věřit? :-)*

**V případě zájmu
se registrujte** přes tento
registrační formulář.

Na co se můžete **těšit?**



**Na diskuzi s naším odborníkem
Tomášem Rosou**, networking s lidmi
z Raiffky a s dalšími studenty.
A možná i na zásadní inspiraci
pro svou kariéru.



Na seznámení s Raiffkou zevnitř,
které se naskytne málokomu,
návštěvu prostor v nejvyšší
kancelářské budově Prahy a pohled
na metropoli jako z letadla.

