



Soumrak lidstva

Může být umělá inteligence nebezpečná pro lidstvo?

Karla Čapka doufám znáte. Pro ty, kdo náhodou neznají jeho slavné drama R. U. R., píše zde o budoucnosti lidstva, která není úplně růžová. Roboti ovládli lidstvo a z lidí se stali jen obyčejní posluhovači, kteří roboty opravují. Původním záměrem přitom bylo, aby si lidstvo žilo v blahobytu, jenže roboti se vzbouřili a lidstvu vyhlásili válku.

V tomto dramatu také došlo k zajímavé zápletce, kde se do sebe dva roboti zamilovali a projevili si city. Dnešní svět je plný digitálních technologií a často slyšíme i o umělé inteligenci. Je však možné, aby se roboti vzbouřili, anebo se do sebe zamilovali a měli city, právě jako v knize Karla Čapka? Na to si v tomto textu odpovíme, a v literatuře ještě na chvíli zůstaneme.

Jak zní zákony robotiky?

Další sci-fi spisovatel, který se podobným tématy zabýval, se jmenoval Isaac Asimov. Ten je známý především proto, že zformuloval tzv. tři zákony robotiky, které zní následovně: Robot nesmí za žádných okolností ublížit člověku. Robot musí poslouchat člověka, pokud to není v rozporu s bodem 1. Robot musí chránit sám sebe, pokud to není v rozporu s bodem 1 a 2. Robotem zde přitom rozumíme zařízení, které má schopnost se rozhodovat a tvořit vlastní úsudek.

Text: Soumrak lidstva

Text připravili: David Kudrna & Petra Sobková,
oddělení vzdělávání NÚKIB



Co je to umělá inteligence?

Vědci dodnes diskutují o tom, co umělá inteligence vlastně je, a co už není. Asi nejčastějším závěrem bývá, že o umělé inteligenci mluvíme tehdy, pokud se umí sama učit, na základě toho rozhodovat, dosahovat cíle a řešit různé typy úkolů, které spolu nemají nic společného.

Dalším problémem je, že se pojmem umělá inteligence pyšní řada zařízení, která na to tak úplně nemají právo. Často se jedná o triky marketingu, protože slovní spojení „umělá inteligence“ je jednoduše atraktivní. Možná jste také slyšeli o tom, že robot pokořil člověka ve hře nebo soutěži a tak dále. Ale takový robot umí v podstatě dokonale jen tuto jednu hru a nic dalšího. Je tedy inteligentní? Není, je naučený.

Jak se umělá inteligence učí návyky a hodnoty od lidí?

I když by si lidé přáli, aby zmíněné literární zákony robotiky fungovaly při skutečném vývoji robotů a umělé inteligence, jejich praktické využití značně pokulhává, což si ukážeme na řadě příkladů. Faktem je, že umělou inteligenci vyvíjí lidé, kteří do vývoje nevědomě přinášejí své návyky a hodnoty. To si pak robot osvojí a aplikuje. Například společnost Amazon vyvinula program pro výběr nových zaměstnanců, kterým chtěla usnadnit práci personalistům při výběru nejlepšího kandidáta.



Tento program nechali učit se z dat o přijatých zaměstnancích za posledních deset let. Jenže postupem času se ukázalo, že tento program některé životopisy znevýhodňuje jiné přímo vyřazuje. Najednou dělalo velký rozdíl pohlaví, barva pleti nebo další podobně faktory. Program vzal data, která analyzoval a neměl potřebu o nich přemýšlet nebo výsledky měnit, bral to jako trend, se kterým dále pracoval. Osvojil si tedy chování personalistů.

Na podobné bariéry narazily i programy, které zkoušely experimentovat v soudnictví nebo policejní praxi. Podobně neúspěšně skončil chatbot, se kterým experimentovala sociální síť Twitter. Chatbot měl za úkol diskutovat s uživateli, a učit se od nich. Jenže se ukázalo, že komunikuje tak, jak komunikují uživatelé s ním. Uměl být rasistický, agresivní, uměl urážet nebo používat sarkasmus. Nevěděl přitom, že jeho chování není vždy správné, protože jen kopíroval jednání lidí.

Problémy při vývoji autonomních automobilů

Další příklad, proč je to s umělou inteligencí složité, mohou být autonomní automobily. Existuje několik úrovní autonomního řízení, přičemž ta nejvyšší by měla být schopná díky velkému množství senzorů a čidel řídit v podstatě bez zásahu řidiče. Otázka ale zní: nakolik se rozhodují samy a do jaké míry jsou ovlivněny lidskými vlastnostmi a morálkou? Při jejich vývoji se například zkoumalo, koho by měl autonomní automobil srazit, pokud by si musel vybrat třeba mezi maminkou s kočárkem a seniorem o hůlce. Mezi obézním člověkem dítětem.

Text: Soumrak lidstva

Text připravili: David Kudrna & Petra Sobková,
oddělení vzdělávání NÚKIB



V průzkumu byly také zajímavé rozdíly mezi lidmi z různých částí světa. Některé kultury totiž pečují o seniory mnohem více než jsme my zvyklí, proto někde upřednostní spíše sražení dítěte. Stavíme tak zařízení, která disponují umělou inteligencí, do situací, které nedokážou řešit ani lidé. K dopravním nehodám totiž dochází a docházet bude, i s autonomním řízením. Tím pádem vyvstává další otázka: Kdo je za nehodu autonomního automobilu vlastně zodpovědný? Automobil? Řidič neřidič? Nebo výrobce?

Umělá inteligence se vyvíjí i bez zásahu člověka

Na druhou stranu jsou i případy, kdy se na výrobě nových technologií a programů podílí více stroje než lidé, aby do nich méně promítali svoje postoje a názory. Vznikají i programy, které se v podstatě napsaly samy a bez zásahu člověka. V takových programech bývají i pasáže, které programátoři nechápou, ale program je pro sebe vyhodnotil jako vhodné a užitečné.

Google překladač se zase sám naučil, že může zvýšit přesnost překladů, pokud využije mezipřeklad přes nějaký třetí jazyk. Jazykovědcům však nedá spát, proč si překladač vybírá pro mezipřeklad jazyky, které spolu nemají nic společného, a přesto je překlad zlepšený. Facebook zase nechal komunikovat dvě umělé inteligence, které využívaly angličtinu. Zjistilo se, že proměnily význam slov a změnily gramatiku, takže nakonec nebylo možné jejich komunikaci vůbec chápat.



Rada na závěr

Ukázali jsme si, že digitální technologie a umělá inteligence mohou být ovlivněny lidským chováním asi víc, než jsme ochotni si připustit. Názor Karla Čapka na tyto a podobné kauzy by mohl být více než zajímavý. Zda nastane apokalypsa a robotický soumrak lidstva tak paradoxně záleží spíše na lidech. Stejně jako v R. U. R. Každopádně každé netvor ví, že záleží na tom, kdo technologie vyrábí a komu se pak nakonec dostanou do rukou.