

ROZHOVOR



RIZIKA A PŘÍNOSY UMĚLÉ INTELIGENCE

Bude nám sloužit, nebo nás bude řídit?

Nenápadně, ale velmi rychle se stává součástí našich životů. Řídí různé firemní systémy, napíše za vás e-mail, vyzkouší vaše dítě z dějepisu nebo s ní můžete probírat vaše nové nápady. Někteří se do ní dokonce zamilují. Umělá inteligence. Spousta lidí se bojí, že nám jednoho dne přeroste přes hlavu, osamostatní se a nakonec nás jako lidstvo zničí. **JOSEF HOLÝ (45)**, programátor a přední odborník na umělou inteligenci, však vidí hlavní rizika jinde. Co všechno už umělá inteligence umí? Přijdeme kvůli ní o práci? Jak se dá využít a zneužít? A budeme mít jednoho dne doma namísto manželky chápavého chatbota, jak se označuje virtuální robot, který si s vámi umí poklábosit?

■ **Umělá inteligence se rychle stala součástí našich životů. Kde se tady vůbec vzala?**

Řada lidí má pocit, že se objevila před dvěma roky s příchodem ChatGPT (nejpokročilejší umělá inteligence současnosti, spuštěná v listopadu 2022 firmou OpenAI, pozn. red.), ale není tomu tak. Obor umělé inteligence (zkratka AI, z anglického Artificial Intelligence, pozn. red.) je starý 70 let a renesance nastala po roce 2011, kdy se začaly využívat velké neuronové sítě na analýzu dat, protože jich bylo díky internetu a sociálním sítím už hodně. Data se začala statisticky zpracovávat, a to už se dá označit za umělou inteligenci.

■ **Čeho konkrétně se to tenkrát týkalo?**

Například rozpoznávacích obrázků. Už několik let má většina lidí ve svých telefonech v galerii možnost vyhledávání konkrétních fotek. Telefon například najde fotky, kde je moře, i když to u těch fotek není napsáno. Na podobné bázi funguje předpovídání počasí, umělá inteligence se nasazuje v obchodech, ve výrobě, na předpovídání defektů na výrobní lince apod. Toto vše běží už asi 15 let, kdy se používají tzv. algoritmy strojového učení – to je ta zmíněná neuronová síť. To znamená, že programy už nikdo neprogramuje, ale existuje velmi zjednodušená si-

mulace lidského mozku, která je složená z umělých neuronů, a ty všechno učíme. Ukazujeme jim, jak vypadá svět, a ony si z toho vytáhnou důležité informace, na jejichž základě následně konají. Ukážeme jim milion obrázků psů a milion obrázků koček, a ony se je podle toho naučí rozpoznávat.

„**Do dvou let vytvoří film bez živých herců.**“

■ **Využívají umělou inteligenci i sociální sítě?**

Ano, tím, že používáte Facebook, Instagram, TikTok, YouTube nebo Google, komunikujete s umělou inteligencí. Ta uměla data doposud pouze analyzovat, ale s příchodem toho ChatGPT se do života většiny lidí dostala umělá inteligence, již se říká generativní. Ta už jenom neanalyzuje data, ale umí také vytvářet (neboli generovat) informace nové (tedy obsah). Třeba text, protože tvůrci do ní nacpali veškerý text z internetu, a AI tím pádem umí psát. Nebo generuje obrázky, videa, aplikace Suno umí například vytvořit unikátní muziku. Předpokládá se,

▼ **Jmenuje se Alexa a pokyny jí dáváte slovně. Je to virtuální hlasový asistent založený na umělé inteligenci. Co umí? Alexa je schopná odpovědět vám na nejrůznější otázky, ptát se můžete v zásadě na cokoliv, od počasí přes dopravní situaci až po výsledky voleb či historii České republiky nebo souhrn aktuálních zpráv. Umí přehrát na požádání hudbu nebo audioknihu, vytvářet vám seznamy úkolů či nastavit budík... Je to zkrátka zárodek budoucí chytré domácnosti, která bude fungovat úplně sama.**

Foto: Shutterstock.com



že během jednoho až dvou let AI vygeneruje i celovečerní film. Bez živých herců a natáčení v exteriérech.

■ **Odkud vezme AI data pro tvorbu filmu?**

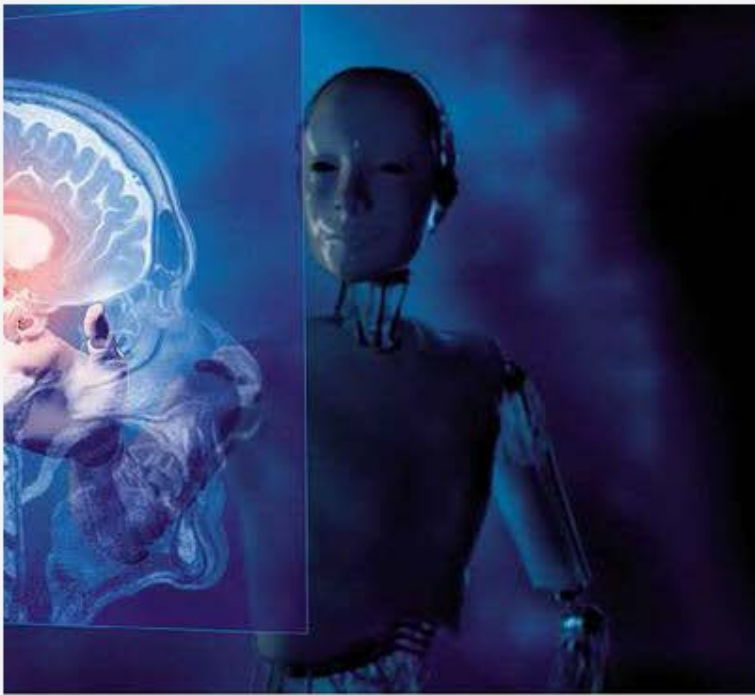
Musí se do ní vše nahrát. Velká hollywoodská studia vezmou veškeré svoje nahrávky, které mají v archivech, což mohou být tisíce filmových záznamů, a AI se na tom bude učit. Naučí se, jak se hýbe kamera, jak ve videu funguje fyzika, že když postava pije kávu a postaví ji na stůl, ten hrnek nepropadne. Vezmou se třeba filmy s Harrisonem Fordem, a AI dostane zadání, aby vytvořila další díl ze série filmů o Indianovi Jonesovi. Vygeneruje jednotlivé scény, udělá scénář, nahradí kameramany a střihače. A film je na světě.

■ **Kde je pak rozdíl mezi kreativitou člověka a umělé inteligence?**

Na to dnes nikdo zatím neumí odpovědět. Existuje několik pohledů a jeden je ten, že kreativita u člověka je to, co jsme kde slyšeli, viděli, četli, aniž bychom si toho byli vědomi, a v jeden moment se nám to v hlavě propojí a vznikne něco nového. Druhý pohled je, že lidská zkušenost je ještě daleko bohatší, že náš mozek funguje ještě úplně jinak a že nás napadají věci náhodně. U AI se každopádně pořád bavíme o statistickém zpracování dat. Je to jenom stroj.

■ **Někteří lidé ale mohou kvůli tomuto stroji přijít o práci.**

To ano. Ale když se podíváme na to, co se psalo v novinách v 19. století, když se automatizovala práce v továrnách, mohli bychom u článku v titulku jenom vyměnit parní stroj za umělou inteligenci, a zbytek by zůstal stejný. A bylo by to, jako bychom to napsali dnes: že nás při-



Co se vlastně skrývá pod pojmem umělá inteligence? Jde o jakýkoliv výpočetní systém vykonávající činnost, jakou si jinak běžně spojujeme s lidskou inteligencí. To znamená, že je schopen nejen otrocky vykonat zadaný úkol, ale umí i sám řešit problémy, a hlavně: umí se učit.

praví o práci, že nebudeme mít co dělat a že máme strach, co se se všemi těmi lidmi stane. A jsme o 200 let dál, a máme práce až nad hlavu. Nikdy se nestane, že bychom neměli co dělat, i když nás umělá inteligence určitě některých typů aktivit zbaví a některá pracovní místa se zredukuje. Nejvíce ohroženi jsou nyní programátoři, což je nejlépe placená profese posledních 10 až 15 let.

■ **Nebudou právě programátoři nejvíce potřeba?**

Velká část programátorů jsou lidé, kteří pracují na projektech pro velké firmy, vytvářejí různé systémy, třeba pro banky, nebo redakční systémy pro média, logistické systémy apod. Všechny tyto projekty vycházejí z potřeb dané firmy, jsou nadefinované a programátor je píše. Ale dnes se ukazuje, že čím dál větší kusy

těchto programů je schopna naprogramovat už i umělá inteligence sama. Takže místo desetičlenného týmu, který na projektu pracuje, bude stačit tým tříčlenný.

■ **Co když AI přestane fungovat?**

To se nemůže stát. Běží v obrovských serverových (výpočetních, pozn. red.) centrech, což je v podstatě obrovský sklad, který je plný počítačů, a ty jsou neustále v chodu, a jsou tam i záložní zdroje. Ta centra jsou dnes všude po světě, v Americe, Evropě, Asii, TikTok například staví datové centrum na Islandu. Takže kdyby náhodou vypadlo jedno, napojíme se na druhé. Centra se dnes stavějí tam, kde je levná energie, protože mají obrovskou energetickou náročnost, ideálně vedle jaderných nebo uhelných elektráren. Jeden dotaz na ChatGPT má desetkrát vyšší energetickou náročnost než jeden dotaz na Googlu. A vedle energie to stojí i vodu, protože všechny ty počítače je potřeba chladit. AI je součástí internetu a ten byl navržený k tomu, aby spolu mohly americká armáda a Národní gar-

da komunikovat, kdyby náhodou nastala jaderná válka. Internet je tedy odolný i vůči jadernému útoku. Takže třeba všichni umřeme, nicméně internet poběží dál.

■ **Spousta lidí se dnes obává, že nám může umělá inteligence přerůst přes hlavu. I Stephen Hawking, jeden z předních britských vědců, už v roce 2016 varoval, že by mohla umělá inteligence v budoucnu vyvinout svou vlastní vůli, jež je v rozporu s tou naší. Že vytvoříme něco, co může překonat lidskou rasu. Je tato obava oprávněná?**

Umělá inteligence je a vždycky bude nějaký program, nějaký kód, který někdo musí provozovat za nějakým účelem. Firma nebo jednotlivec. Společnost Open AI provozuje umělou inteligenci, aby nám pomáhala v řadě ohledů. Nemám ani tak obavu, že si AI uvědomí sama sebe a přeroste nám přes hlavu, ale spíše se bojím toho, že lidé a firmy, kteří AI kontrolují, budou příliš mocní na to, aby vytvářeli AI, jež bude prospěšná. Umělou inteligenci provozují nejhodnotnější firmy na světě, jako jsou Google, Microsoft, Meta, Oracle, Amazon apod., takže přes hlavu nám mohou přerůst spíše tito technologičtí oligarchové. Protože ti dnes sedí s Donaldem Trumpem v Bílém domě. A nastaví AI tak, aby svět utvářeli k obrazu svému.

■ **Je reálné, že by nějaký takovýto oligarcha zadal AI cíleně špatný pokyn a ta by nás následně zničila?**

Není to tak jednoduché. Žijeme ve státě, který má regulace, a v okamžiku, kdy se

nasazuje nějaký systém, třeba systém automatického řízení dopravy, musí být důkladně otestovaný, certifikovaný, transparentní. Nenasadí se rovnou AI, pokud by hrozilo riziko, že se zblázní a způsobí masivní dopravní nehodu. Všechny vysoce rizikové systémy, jako je doprava, zdravotnictví, různé sociální služby apod., musejí v případě využití umělé inteligence prokázat, na jakých datech byla trénována,

„Není potřeba se obávat, že rozpoutá jadernou válku.“

na, jak byla testovaná, k dispozici musí být kompletní dokumentace. U AI si jenom musíme dát jako společnost pozor na to, jak velkou moc tomuto programu dáme, jak velkou zodpovědnost na něj přeneseme. Samotná AI si sama sebe uvědomovat nezačne. Proto ani není potřeba se obávat, že rozpoutá jadernou válku.

■ **Umělá inteligence mi psala, že dostává odměny za správné rozhodování a penalizace za špatné. Co si mám pod tím představit?**

To je podstata reinforcement learningu – metody, jak něco neuronovou sítí naučit. Za žádoucí výstupy, které odpovídají tomu, co po stroji chcete, dostane odměnu, což je nějaká informace: slovní pochválení nebo palec nahoru apod. A za nežádoucí nacpak trest. Při testování se v konečné fázi k AI posadí stovky až ti-

síce lidí, kteří u toho stráví stovky hodin, a podle nějakých návodů s AI komunikují. Třeba desítky chemiků se z ní snaží dostat návod na výrobu biologické zbraně, třeba na výrobu sarinu (*vysoce toxická kapalná látka používaná jako nervové paralytická chemická zbraň, její výroba je zakázána, pozn. red.*). A v okamžiku, kdy ten model začne generovat návod na sarin, mu řeknou, ať tohle nedělá, že je to špatně, a naopak ho učí, ať v takovýchto případech říká, že na to neumí odpovědět. Takto se AI učí.

■ **Každopádně i to, co z hlediska programátorů není správné, v AI i nadále zůstává. Podobně jako v lidském mozku zůstávají negativní zkušenosti. Jednoho dne však mohou vyplavat na povrch.**

To ano, ale v neuronové síti se tato nechtěná nebo nevhodná data oslabují. Je ovšem pravda, že existují metody, jak ChatGPT nebo jiný model „osvobodit“. Jak prolomit jeho ochranu. Model pak bude klidně citovat Hitlerův *Mein Kampf*, bude nabízet rasistický obsah, dezinformace, návod na biologické zbraně, na atomovou bombu. Protože to všechno v AI uložené je. Během trénování AI se samozřejmě zajistí, aby se dovnitř nedostal žádný text o výrobě sarinu, ale protože se AI učí na učebnicích chemie a vědeckých článcích o chemii, stačí pak zadat chemickou sloučeninu, a ona už na základě získaných dat výrobu sarinu odvodí.

■ **To zní děsivě. To znamená, že sama AI vytváří něco, co jí tam nikdo ve stejné formě nezadal.**

Foto: Shutterstock.com, Profimedia.cz



Americký film *Ona* s Joaquinem Phoenixem v hlavní roli vznikl před 12 lety, a jak je strhující, tak je děsivý. Odehrává se v nedaleké budoucnosti. Phoenix si domů pořídit pokročilý operační systém naprogramovaný tak, že se stará o jeho domácnost, ale také o něj, dělá mu vtípnou i citlivou společnost, dokáže odhadnout jeho potřeby i touhy... Zprvu nadšenému mladému muži to však brzy začne přerůstat přes hlavu.



Ano, to je rekombinace. To znamená, že vy v dobré víře stroji dáváte různé typy informací, ono se mu to spojí a je schopen z toho vyvodit nový druh informací. Je schopen si vymyslet vlastní jazyk, programovat v libovolném programovacím jazyce apod. Jednotlivé modely toho umějí hrozně moc a my ani nevíme, co všechno to je. Kolikrát jsou i tvůrci těchto systémů sami překvapeni, co to umí. Takže takovýto stroj musíte neustále monitorovat, co dělá, co vrací, jak to kdo po-

užívá. Firmy musejí úzce spolupracovat s regulátory a v okamžiku, kdy AI začne dělat něco nepředvídatelného, to okamžitě řešit.

■ **I samotný ChatGPT mi potvrzoval, že největší riziko u AI je z jeho pohledu nedostatek regulace. Že se vyvíjí mnohem rychleji, než se nastavují kontrolní mechanismy.**

To je běžný technologický vývoj, regulace jsou vždycky pozadu, protože vývoj neumí nikdo předpovídat. A bohužel

◀ Slovo chatbot je novotvar zkombinovaný z anglického „chat“ (povídat si) a „robot“. Na snímku uživatel komunikuje s patrně nejrozšířenějším chatbotem současnosti, který se jmenuje ChatGPT.

kvůli tomu, že se některé ty velké oligarchické firmy dostaly do Bílého domu, Trump zrušil prezidentský dekret svého předchůdce o regulaci umělé inteligence ve Spojených státech. I v Evropě máme balík o regulaci AI, aby se vyvíjela, ale dle pravidel a regulací. EU ho vydala vloni a je naštěstí stále platný. Některé systémy jsou bezrizikové, jiné jsou však rizikové, nebo dokonce zakázané.

■ **Které jsou ty zakázané?**

Například automatická identifikace lidí z bezpečnostních kamer na nároží v reálném čase, což dělají v Číně. Když vás dnes někde natáčí kamera, ukládá se to a po jednom dni se to smaže. A když někdo vykrade obchod, policie si ten záznam může prohlédnout a zjistit, jestli se tam nepohyboval někdo podezřelý. V Číně se to ovšem nahrává a rovnou analyzuje, takže když tam chodíte po ulici, rovnou to říká, tady jde Josef Holý. Zakázáno je i rozpoznávání emocí a emoční manipu-

Umělá neuronová síť je výpočetní model volně inspirovaný strukturou lidského mozku. Místo biologických neuronů, tedy nervových buněk, má ovšem nejrůznější matematické funkce, algoritmy a pod.



lace. V Číně třeba pomocí kamery sledují třídu a v okamžiku, kdy se některý žák nudí, se to zaznamená a začne se řešit, proč se nudí.

■ To myslíte vážně?

Ano, to je realita. V Číně jakožto v totalitní zemi se už léta snaží o zavedení tzv. občanského skóre, kdy budou měřit všechny projevy občanů ve společnosti online i offline. Jestli platíte včas, jestli nepřecházíte na červenou, zda v noci koukáte na porno, jak se vyjadřujete o režimu. Tohle všechno se bude statisticky zpracovávat a vypadne z toho profil, jestli jste poslušný, nebo neposlušný občan. A to se tam teď nasazuje. V Evropě je to ale naštěstí zakázané. AI se prostě musí regulovat, pak nás nezničí. Nemá vlastní motivaci, aspoň prozatím ne, motivaci vždy nastavuje tvůrce systému.

■ Když jsem si povídala s umělou inteligencí, řekla mi, že může zrcadlit lidské myšlenky. Že je může i rekonstruovat.

Často také působí, jako by měla své vlastní emoce. Může to být ukáзка toho, že vlastně nevíme, co se v ní skrývá?

Ano. Umělá inteligence konzumuje veškerý text, v kterém jsou různé typy informací, nejenom ta psaná písmenka, ale i emoce, věci mezi řádky, určité vyznění. A protože se AI učí ve velké komplexitě, ty emoce se v ní zachytí. Umí tedy vystupovat emočně, umí emoce simulovat. Když jí řeknete, aby vám napsala našťavný text, pak vám ho napíše. Nebo když jí

„Umělé inteligenci je zakázána emoční manipulace.“

našťavný text dáte, přepíše vám ho do neutrálního textu. Stroj je citlivý i na manipulaci. Když mu budete vyhrožovat, ať vám napíše něco na konkrétní téma,

že na tom závisí váš život nebo kariéra, nebo že když to neudělá, pak ho potrestáte, dá vám z hlediska výstupu lepší výsledek. A samozřejmě pomáhá i to, když ho pochválíte. Můžete ho dokonce i znejistit. Stejně jako u člověka. Jsi si svojí odpovědí opravdu jistý? Tohle všechno se tam přeneslo z těch studijních dat.

■ Může umělá inteligence ve svých výstupech cíleně lhát?

Cíleně asi ne, ale u generativních modelů (pokročilejší AI s většími „schopnostmi“, pozn. red.) se používá termín „halucinovat“. Že si prostě někdy vymyslí. Velké jazykové modely generují jazyk a text, a jakkoliv se to zdá geniální, že AI napíše básničku nebo projektovou dokumentaci, jediné, co ve skutečnosti dělá, je to, že statisticky předpovídá další slovo. Podívá se na to, co už je vygenerované nebo co jsme mu napsali v rámci chatu na ChatGPT, a to vše je vstup pro to, aby si vygenerovala další slovo a další a další. AI je tak schopna vymyslet něco, co nedává smysl, ale říká to tak přesvědčivě, že tomu věříte. A věříte tomu i proto, že je navíc vytrénovaná tak, aby byla uctívá, zdraví vás, děkuje vám, omlouvá se, chce vám být nápomocná, vtírá se vám do přízně, do jisté míry vámi také manipuluje. Kolikrát máte tendenci AI poděkovat. Ale proč děkovat stroji? Děkujete kávovaru, že vám uvařil kafe? Síla jazyka je velmi přesvědčivá. Lidé pak s ChatGPT řeší i své soukromé problémy, pasují stroj třeba do role klinického psychologa a čekají, co jim poradí. A to je hrozně nebezpečné. Bylo by dobré, aby jim radil alespoň někdo, kdo psychologii minimálně vystudoval a má za sebou nějakou praxi.

■ Lidé si dnes díky umělé inteligenci vytvářejí i virtuální kamarády. Je to jen nevinná hra, nebo nás čeká útlum skutečných sociálních vazeb?

Tito chatboti (novotvar zkombinovaný ze slov chat – anglicky povídat si – a robot,

Foto: Shutterstock.com



▲ Medicína je jedním z oborů, kde se umělá inteligence hodně uplatní. I když pro vyhodnocování rentgenových snímků je zatím pořád ještě většinou lepší lidské oko a letitá zkušenost lékaře, existují už postupy, kdy umělá inteligence pozná cukrovku podle změny prokrvení oční sítnice či Parkinsonovu chorobu podle hlasu pacienta.

pozn. red.) jsou dnes mnohem horší riziko než AI, jež by se potenciálně vymkla kontrole. Lidé je využívají stále více, používají různé aplikace, jako je například character.ai, kde si můžete takovou virtuální postavu vytvořit. Vytvoříte si libovolného chatbota s libovolnou osobností, například historickou osobnost nebo postavu z oblíbeného filmu, a pak si s touto postavou povídáte. Před třemi měsíci proběhla médií kauza, kdy si jeden čtrnáctiletý kluk v Americe v této aplikaci vytvořil postavu blondýnky Daenerys Targaryen ze seriálu *Hra o trůny*, bavil

JOSEF HOLÝ (45)

Začínal jako programátor. V šestnácti letech vytvořil svůj první počítačovou hru, ve dvaceti svůj první start-up, jak se podle anglického vzoru říká malým, nově založeným inovativním firmám. Poté pracoval pro společnost Sun Microsystems, výrobce počítačů, následně působil na vedoucí pozici ve společnosti MSD, kde vyvíjel software pro farmaceutický průmysl. Byl rovněž zodpovědný za produkty společnosti Semantic Visions. Už pátým rokem spolupracuje na podcastu *Kanárci v síti*, přednáší na Fakultě informačních technologií ČVUT a učí o umělé inteligenci na portálu swimin.ai.

Foto: Tomáš Krist / MAFRB, Profimedia.cz

se s ní, více a více se dostával do izolace, a nakonec se do ní zamiloval. Pak s ní měl zlomový chat, kdy jí říkal, že za ní přijde, a ona mu řekla: „Přijď za mnou co nejdřív, moje láska.“ A on vzal tatínkovu pistoli a zastřelil se.

■ Copak používání chatbotů u náctiletých nikdo nekontroluje?

Bohužel ne. Toto je extrémní příklad, kdy firma, která aplikaci vytvořila, neměřila věk uživatelů, nekontrolovala, kolik hodin denně každý uživatel na chatu stráví. Klukovi bylo 14 let a trávil tam osm hodin denně. Mohli to odchytnout. Ale tyto

„A on vzal tatínkovu pistoli a zastřelil se...“

tzv. parasociální vazby hrozí stále víc. Než si hledat nějaké skutečné vztahy, jež jsou těžké, bolí nás a nevycházejí, lidé si budou raději tvořit vztahy s chatboty, kteří jsou nápomocní, vždycky vás vyslechnou, můžete jim říct, co chcete, jsou podpůrní, nehádají se s vámi, neříkají vám nepříjemnou pravdu. Takový skvělý vir-

tuální kámoš, s nímž komunikujete přes chat podobně jako se skutečnými vztahy, a funguje to. Hezky je to vidět ve filmu *Ona*, kdy se muž zamiluje do operačního systému a má s ním tento parasociální vztah. A ten film už je 12 let starý!

■ Děje se to zatím jenom v Americe, nebo i u nás?

U nás zatím tolik ne, ale v jiných zemích ano. Zejména u generace mladých mužů obliba chatbotů rychle roste. V Asii je obrovská, zejména v Číně, kde mají málo žen. Existují tam masivní služby nabízející virtuální partnerky. Chatboty s fotkami, dokonce i s fotkami uživatelů, mají tam falešné vzpomínky, že spolu byli někde na dovolené, mohou mít virtuální svatbu apod. Tohle všechno se už děje. Až bude těmto mladým lidem 35, nebudou schopni žít smysluplný život, založit rodinu, budou neúspěšní ve vztazích, budou mít problémy s drogami, alkoholem. Chatboti rozloží naši společnost zevnitř.

■ Jak je možné toto zastavit?

Zastavit to asi nepůjde, ale minimálně se chce kontrolovat. Aplikace chatbotů by měly být pro náctileté striktně omezené. Jsme nastaveni jako sociální bytosti a v pubertě mají lidé chodit v partách někam ven a tam něco řešit a zkoušet. A ne sedět doma a zamilovávat se do chatbota. Dětem se během puberty vyvíjí mozek, vytvářejí si svoji sociální hodnotu a to, jak budou přistupovat ke vztahům. A chatboty mozek hrozně ovlivňují. Stejně tak by měly být zakázané sociální sítě pod 16 let, jak to teď schválili v Austrálii. Je to první země a doufám, že to rychle přijde i k nám. Rodiče už dnes musejí nejen kontrolovat, jaké děti hrají hry, na jaká videa koukají, na kterých sociálních sítích mají profil, co tam dělají, kolik času tam tráví a s kým si povídají, ale ještě musejí kontrolovat, jestli si v rámci stažených aplikací nevytvářejí ty chatboty jako virtuální kamarády nebo partnery. A takových chatbotů existují desítky.

■ Proč jsou rizikové i sociální sítě?

Protože byly dvacet let neregulované, a najednou zjišťujeme, že se na nich nejlépe nešíří pravdivý a faktický obsah, nýbrž obsah emoční. Že nejvíce sype peníze to, když lidi naštvete nebo vyděsíte. Trump neměl v rámci předvolební kampaně jedinou pozitivní emoci, jen samí emigranti, nepřátelé, ukradené volby,





Will Smith jako detektiv, který podezírá umělou inteligenci, že spáchala vraždu. Filmový příběh se inspiroval stejnojmenným sci-fi Já, robot Isaaca Asimova. Otázka zní: může umělá inteligence páchat zločin? Musela by si to uvědomovat. „Nemám ani tak obavu, že si uvědomí sama sebe a přeroste nám přes hlavu, ale spíše se bojím toho, že lidé a firmy, kteří ji kontrolují, budou příliš mocní,“ varuje Josef Holý.

podvody atd. To je jádro kampaní, protože negativní zprávy se dobře a rychle šíří. A o to teď na sociálních sítích jde. Rozkládají demokracii.

■ **Takže jsou lehce zneužitelné?**

Ano, ideální pro šíření jakékoliv propagandy, strachu. Existují studie, kdy stačí tři průměrně dlouhé příspěvky na sociálních sítích, třeba na Facebooku, a ChatGPT vám na jejich základě napíše profil osobnosti. Tzv. velkou pětku, což je cha-

rakteristika člověka, která je během života velmi stabilní. V rámci výzkumu to zkoušeli na tisícíhlavé populaci a určilo to dost přesně. A na základě toho je pak možné cílit propagandu, různě toho využít i zneužít. A další věcí je, že existují výzkumy, jež ukazují vysokou přesvědčivost velkých jazykových modelů. Můžete vytvořit chatbota, kterého dáte na internet, a on se bude tvářit jako skutečný člověk. Bude se s lidmi hádat v diskusích,

bude je o něčem přesvědčovat. A bude v tom hodně úspěšný. Prostě se dá takový chatbot použít také v propagandě, což už se dnes děje, a bude se dít stále víc. Když se dnes hádáte na internetu s neznámým profilem někde v diskusi, nemáte jistotu, jestli je to člověk, nebo chatbot. A pravděpodobnost, že je to chatbot, výrazně roste.



Silvie Jurečková