

Inteligentní infrastruktury: požehnání, nebo prokletí?

Jiří Wiedermann

Centrum Karla Čapka pro výzkum hodnot ve vědě a technice
Ústav informatiky AV ČR
Pod Vodárenskou věží 2, 182 07 Praha, Česká republika
Email: jiri.wiedermann@cs.cas.cz

Abstrakt

Dnešní doba neobyčejně přeje rozvoji tzv. specializované umělé inteligence, která je schopna řešit úzce vymezené problémy v pečlivě konstruovaných omezených prostředích. Proto se začínají uvažovat rozsáhlé, tzv. inteligentní infrastruktury, které využívají řadu spolupracujících specializovaných umělých inteligencí, jako jsou marketingové systémy, doporučovací systémy a sociální kreditní systémy, jdoucí přes různé oblasti života lidí a společnosti. Jejich cílem je učinit poskytování různých služeb efektivnější, operativnější, bezpečnější a spravedlivější. V příspěvku nastíníme architekturu těchto systémů a naznačíme výzkumné problémy spjaté s jejich vývojem. Společenské přijetí zejména hodnotících systémů je kontroverzní záležitost vedoucí k řadě legálních a etických problémů.

1 Úvod

Současný vývoj v umělé inteligenci narazil do zdi (Buchanan, 2018). Experti přestávají pracovat na vývoji umělé inteligence podobné lidské inteligenci (LeVine, 2018). Budoucnost umělé inteligence není o počítači, který myslí tak, jako člověk (Sandhu, 2018). To vše jsou titulky zpráv z posledních měsíců, které referují o výročí významných expertů o současném stavu umělé inteligence. Pokud budoucnost umělé inteligence není o počítači, který myslí tak, jako člověk, tak o čem je ta budoucnost?

Jednou z cest, jak přemýšlet o budoucnosti umělé inteligence, je slevit z podmínky o počítači, který myslí tak jako člověk. Rozumnou cestou se zdá být využití různých existujících specializovaných inteligencí, které mohou spolupracovat na řešení stejných problémů viděných z různých aspektů a které mohou využívat pro svoji činnost stejnou technologickou infrastrukturu. To je základ myšlenky inteligentních infrastruktur. Cílem budování inteligentních infrastruktur je učinit poskytování různých služeb, jak individuálních, tak i sociálních, efektivnější, operativnější, bezpečnější a spravedlivější.

Aktivním propagátorem myšlenky inteligentních infrastruktur je Michael I. Jordan (2018), americký vědec, profesor na Kalifornské univerzitě v Berkeley,

vedoucí osobnost v oblasti strojového učení. Jeho představa, prezentovaná v poslední době v sérii jeho přednášek na významných konferencích z oboru umělé inteligence vystoupení je ta, která, kdybychom použili metaforu z nadpisu tohoto příspěvku, by znamenala „požehnání“ pro uživatele inteligentní infrastruktury. Profesor Jordan však nezmiňuje donedávna málo známý vývoj kolem tzv. systému sociálních kreditů, což je také inteligentní infrastruktura, která se v současné době realizuje v Číně, u které se většina pozorovatelů z okruhu západní civilizace domnívá, že je spíše „prokletím“ celé společnosti, nežli „požehnáním“ (Botsman, 2017).

Cílem tohoto příspěvku je ukázat, že oba dva přístupy k inteligentním infrastrukturám — jak přístup profesora Jordana, tak i čínský přístup, vycházejí ze stejných premis a využívají stejné informační technologie. Přínosem tohoto článku je jisté zobecnění pojmu inteligentních infrastruktur tak, aby jejich architektura pokrývala oba zmíněné přístupy. Zmíníme i nové výzvy v oblasti výpočetních síťových technologií, v oblasti zpracování velkých dat a v oblasti specializované umělé inteligence, které před nás staví realizace inteligentních infrastruktur. Rozdíl mezi oběma přístupy je hlavně v zamýšleném využití těchto infrastruktur, což vede k řadě etických problémů se závažnými důsledky, se kterými jsme zatím nebyli konfrontováni ani ve své komplexnosti, ani v jejich rozsahu.

2 Vize inteligentních infrastruktur

V současné době neexistuje žádná obecně uznávaná definice pojmu inteligentních infrastruktur. I když je tento pojem ve vývoji, můžeme alespoň uvést vizi, popisující hlavní charakteristiky těchto systémů.

Z pohledu zúčastněných vědních disciplin a jejich poslání jsou inteligentní infrastruktury rozsáhlé výpočetní systémy, které využívají informační technologie specializovaných umělých inteligencí na průniku *informatiky, umělé inteligence, ekonomiky, sociologie a etiky* za účelem efektivního, operativního, bezpečného a spravedlivého poskytování různých služeb.

Činnost inteligentní infrastruktury určují pravidla hry — kdo a jak, za jakých podmínek, se může podílet na činnosti systému, jaké tím získá výhody, a

jaké nevýhody.

Z pohledu jejich návrhu a jejich vnitřní architektury se skládají informační infrastruktury z pěti specifických vrstev, které zhruba odpovídají shora zmíněným pěti vědním disciplínám.

Základem je dnes již do jisté míry „klasická“ infrastruktura internetu věcí (IoT). *Internet věcí* je heterogenní síť propojených počítačů, mobilních telefonů, dopravních prostředků, domácích spotřebičů a jiných objektů, které obsahují elektroniku, software, senzory a aktuátory a mají jednotný identifikátor (většinou ve formě IP adresy) a schopnost navzájem komunikovat prostřednictvím sítě. To všechno jsou prostředky, které využívají zejména nástroje informatiky resp. softwarového inženýrství. Prostřednictvím objektů infrastruktury se stávají účastníky sítě uživatelé těchto objektů.

Dále požadujeme, aby bylo možné využívat data a datové toky pro budování znalostí o objektech, které tvoří síť. Za tím účelem musí být data, plynoucí sítí, koherentní, zapadající do nějaké formální či méně formální *znalostní teorie* a využitelná pro odvozování (inferenci) znalostí v rámci této teorie. Zde přicházejí do hry specializované umělé inteligence. Množné číslo používáme proto, že nad stejnou sítí může fungovat několik znalostních teorií, každá z nich specializovaná na specifický typ dat o objektech a budující specifickou znalostní bázi. Např. jedna báze může být zaměřena na zdravotní informace, další na dopravní, sociální, atp. Také zde může být mechanismus — *preferenční systém* — určující preference jednotlivých objektů na základě jejich minulého anebo statistického chování v různých sledovaných oblastech zájmu.

Na základě informací ze znalostní báze o každém objektu a jeho interakcích a minulých akcích může dále *tržní mechanismus* společně se znalostí preference vytvářet pro daný objekt *doporučení* pro vykonání nějaké nové akce nebo pro realizaci nějakého rozhodnutí. Na této úrovni již inteligentní infrastruktura funguje jako inteligentní trh (či tržiště), v klasickém ekonomickém smyslu. Takový systém je v pozadí systému jako je Uber pro dopravní služby, Airbnb pro ubytovací služby, doporučovací systém firmy Amazon pro tipy na nákup preferovaného zboží, atd.

Nad takto fungujícím inteligentní infrastrukturou lze dále implementovat *kreditní systém*, který systematicky sbírá kredity (hodnoty) přiřazené rozhodnutím či akcím realizovanými jednotlivými objekty. Pokud jsou tyto kredity sociologické povahy, vztahující se k rozvoji, struktuře a fungování společnosti, tak výsledný systém tvoří sociální kreditní systém. V závislosti na jejich povaze lze kredity využívat pro hodnocení daného objektu z různých sociologických aspektů: lze např. hodnotit jeho schopnost splácet půjčky, potenciál jeho koupěschopnosti, společenské úspěšnosti, hodnocení zdravotního stavu za účelem zdravotního pojištění, atd. Jiným příkladem kreditního systému je systém používaný ve vysokém školství, který hodnotí úspěšnost

studia.

Smysl kreditnímu systému dává poslední, pátá vrstva inteligentních infrastruktur: *etika*. Je to systém morálních principů, které mají vliv na to, jaké rozhodnutí činí lidé (či obecně, objekty). To samozřejmě závisí na tom, co je obecně považováno za dobré pro jedince nebo pro společnost.

Počet pěti vrstev architektury inteligentních infrastruktur je v jistém smyslu maximalistický — samozřejmě, je možné uvažovat takové systémy, ve kterých některá z vrstev chybí. Pak dostaneme vlastně příklady existujících systémů tohoto druhu, jako jsou např. již zmíněné rezervační systémy typu Uber nebo Airbnb, různé doporučovací systémy anebo bankovní kreditní systémy, elektronická evidence tržeb, elektronický recept, atp.

3 Výzkumné výzvy spjaté s inteligentními infrastrukturami

I z předchozího stručného popisu architektury a činnosti inteligentních infrastruktur je zřejmé, že ve své maximální verzi se jedná pravděpodobně o nejkompaktnější systémy pro zpracování dat, které známe, resp. přesněji: dovedeme si reálně představit (protože zatím, v době psaní tohoto příspěvku, žádný takový systém neexistuje).

Složitost takových systémů je daná několika faktory, a to jak technologickými, tak i bezpečnostními a etickými.

Prvním z nich je rozsah takových systémů. V praxi se může jednat až o systémy obhospodařující až stamilióny subjektů (viz připravovaný čínský systém sociálních kreditů, zmíněný v další části). Z tohoto rozsahu vyplývá i obrovské množství dat („big data“), které musí systém obhospodařovat. Tento úkol se liší od běžného off-line zpracování velkých dat, protože data do systému neustále, on-line způsobem přibývají, čím se mění preference uživatelů, a změny, které se promítají třeba do preferencí a doporučení, je potřebné v reálném čase zpracovat.

Zde nastává problém, kde fyzicky je vhodné držet příslušná data. Pokud jsou příliš vzdálená a centralizovaná v datových centrech (tzv. cloud-based networks, síť datových mraků), narážíme na problém jejich rychlé dostupnosti (to je typický problém IoT). Většina interakcí má totiž charakter, že využívá související data získaná z blízkého okolí a ze stejné sémantické oblasti, a současně využívají vysokorychlostní (např. 5G) síť. V takových případech je potřebné služby datového mraku distribuovat po síti na její „okraje,“ blíže k výpočetním zdrojům, kde probíhají intenzivní výpočty (tzv. cloud-edge computing). V těchto oblastech stále probíhá intenzivní výzkum.

Dalším problémem je vzájemná provázanost rozhodnutí doporučovacích systémů. To nastává v si-

tuacích, kdy jsou k dispozici omezené zdroje vzhledem k počtu zájemců o ně (a to je prakticky vždy), a tudíž nelze všem uživatelům doporučit stejná rozhodnutí, protože zdroje se mohou vyčerpat. Např. v případě dopravní zácpy nelze všechny automobily poslat stejnou objížděnou. Jiný případ nastane, kdy ve stejné situaci (např. „hledám asijskou restauraci v pěší vzdálenosti“) doporučení závisí na mé okamžité poloze. V některých případech dokonce nelze uspokojit všechny zájemce. Tehdy je potřebné volit mezi nimi a vybrat, kdo bude a kdo nebude uspokojen. Zde vstupuje do hry etika ve spolupráci se s kreditním systémem.

Jiný problém s daty je jejich původ. Jsou spolehlivé? Nejsou zastaralá? Učení neuronových sítí konané z nespolehlivých či zastaralých dat zřejmě nemůže být optimální a tedy i rozhodnutí takových sítí budou nespolehlivá.

Úkolem inteligentních infrastruktur je také vydávání různých doporučení. Pokud mají být taková doporučení důvěryhodná, musí být k dispozici vysvětlení, jak a proč k danému závěru systém dospěl. To, jak je známo, je problém, zejména pokud se využívají neuronové sítě, o kterých nevíme, jak ke svým výsledkům dospívají.

Kombinací tržních mechanismů s doporučovacími systémy, pracujícími s velkými daty, plně nerozumíme. Proto je důležitá extrakce příslušných znalostí z dat o práci takových systémů pomocí učení bez učitele. O tom také skoro nic nevíme. Zde hrají důležitou roli kauzální mechanismy. Obecně se má za to, že neuronové sítě nejsou vhodným prostředkem pro zjišťování kauzality. Jak jinak tedy řešit uvedené problémy?

Také nečekané situace jsou problémem. Jak zajistit robustnost a férovost řešení i v takových podmínkách?

Ochrana soukromí, vlastnictví dat a sdílení dat mezi organizacemi a soukromíky je další velký okruh problémů, se kterým v podmínkách inteligentních infrastruktur nemáme zkušenosti. To už jsou vlastně etické problémy, kterým se budeme věnovat v další sekci.

4 Etické problémy inteligentních infrastruktur

Inteligentní infrastruktury vyžadují pro svoji činnost o každém objektu množství dat privátního charakteru. Zde vznikají otázky, jak se taková data získají, kde se uchovávají, komu patří, jak jsou chráněna. To jsou obecné problémy s daty personálního charakteru, které jdou za rozsah i obsah tohoto příspěvku. Dalším problémem je asimilace dat do různých systémů specializovaných inteligencí. Data o daném objektu v obecném případě pocházejí z mnoha zdrojů, soukromých (např. Facebook), veřejných (různé

rejstříky), neveřejných (firemní interní informace o zákaznících, registr trestů, informace od státních organizací, pojišťoven, lékařů, atp.). Kde jsou hranice svobodného nakládání s takovými informacemi? Kdo může mít přístup k nim, ze jakých okolností?

Tyto a podobné problémy ve své obecnosti řeší etika. Problém je samozřejmě v tom, že etika řeší tyto problémy v obecnosti, a většinou se nezabývá konkrétními případy, které jsou ovšem pro inteligentní infrastruktury nejzajímavější. Nejlépe je to vidět na případě sociálních kreditních systémů. Tyto systémy mají dlouhou tradici sahající před věk počítačových informačních technologií. Typickým příkladem je důchodový systém, což je způsob, jakým obyvatelstvo dané země financuje definované důchodové dávky pomocí sběru kreditů během doby své pracovní činnosti. Extrémním případem, kde etika hraje hlavní roli ve využití sociálních kreditních systémů, je v současné době budovaný čínský sociální kreditní systém. Jeho cílem je monitorování, hodnocení a řízení chování tržních účastníků tohoto systému v omnoho širším měřítku, než stávající kreditní systémy. Jestli se experiment podaří, tak systém posílí důmyslným způsobem možnosti čínské vlády řídit a ladit tržní a sociální mechanismy a průmyslovou politiku v celé zemi. Dobrý popis tohoto systému je např. v práci (Meissner, 2017).

Čínský sociální kreditní systém je navržen tak, aby motivoval dodržování právního řádu. Občané získávají body za dobré skutky, jako je např. dobrovolnictví, darování krve, nebo přilákání investic do města. Body lze na druhé straně ztrácet za různé přestupky, jako jsou porušení dopravních pravidel (dokonce i přechod chodců na červenou — zde se využívají on-line kamerové systémy pro rozeznávání osob podle obličeje a chůze), neplacení daní, nestaráni se o své staré rodiče, nekontrolované hraní počítačových her, atd. Získané kredity pak mají vliv na různé benefity: přístup k neplaceným zdravotním službám, slevu na nájemném, výhody v bankovníctví, přístup ke kvalitnímu vzdělání, k rychlému internetu, atd. Málo kreditů znamená ztrátu výše zmíněných benefitů, státní podpory, nemožnost pracovat ve státních službách, omezení cestování letadlem či rychlovlakem, atd.

Z pohledu západní civilizace představuje čínský kreditní systém pro místní režim prostředek poháněný umělou inteligencí, který porušuje lidská práva (viz např. (Thomson, 2017)). Zřídka však čteme, že tento mechanismus vítá většina občanů Číny, kteří mají dost nekontrolovatelných zpronevěr, korupce, podvodných produktů, absence komunitních a státních správních infrastruktur a nefunkčního zdravotnictví. Současně to může být důsledek tisícileté čínské tradice motivovat morální chování lidí a zcela jiného pojetí pojmů soukromí a osobní svobody.

5 Závěr

Naše vize inteligentních infrastruktur představuje pokus o maximální využití současně známých specializovaných umělých inteligencí. Je to způsob, jak generovat do jisté míry inteligentní chování ve formě „užitečných rad“ a informací pomocí výpočetních agentů, ze kterých žádný zvlášť takového chování není schopen. Ve své výsledné podobě se vlastně jedná o interaktivní hry s velkým počtem hráčů, které ovšem neslouží pro zábavu, nýbrž pro usměrňování a motivování smysluplného chování účastníků hry, případně pro naplňování jejich potřeb.

Z etického hlediska takové systémy před nás staví novou výzvu. Doposud jsme totiž ve velké většině vnímali umělou inteligenci jako inteligenci, vlastní nějakému zařízení, třeba robotu či počítači, které funguje do jisté míry autonomně a tudíž se může „vymknout“ z řízení lidí, dokonce může převýšit inteligenci lidí a konat proti jejich zájmům. To není typický případ inteligentních infrastruktur. Ty celkem jistě nejsou autonomní, nekonají nic jiného, nežli sbírají informace, které jim lidé dovolí a dají k dispozici, a podle návodů, dodaných lidmi, poskytují „rady“ a služby, když jsou o to uživatelem požádány. Tyto systémy samozřejmě neví, co činí, a za své konání nemohou být odpovědní. Lidé je mohou, ale nemusí budovat a poslouchat. Dovedeme se takové výzvě postavit? Stanou se inteligentní infrastruktury požehnáním, nebo prokletím?

Poděkování

Tento příspěvek vznikl v rámci Centra Karla Čapka pro výzkum hodnot ve vědě a technice, za částečné podpory institucionálního plánu ÚI AV ČR RVO 67985807 a programu Strategie AV21 „Naděje a rizika digitálního věku“.

Literatura

- Botsman, R. (2017). Big data meets Big Brother as China moves to rate its citizens. In: WIRED, <https://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion>
- Buchanan, M. (2018). Our Robot Overlords Might Be Delayed. Research in artificial intelligence seems to have hit a wall. In: Bloomberg, <https://www.bloomberg.com/opinion/articles/2018-04-04/artificial-intelligence-research-might-have-hit-a-wall>, 2018
- Jordan, M. (2018). Artificial Intelligence - The Revolution Hasn't Happened Yet. In: Medium, April 18, 2018, <https://medium.com/@mijordan3/artificial-intelligence-the-revolution-hasnt-happened-yet-5e1d5812e1e7>

LeVine, S. (2018). AI researchers are halting work on human-like machines. In: AXIOS, <https://www.axios.com/ai-researchers-are-halting-work-on-human-like-machines-c7c821c2-59b1-40da-bdb4-e2a368b215aa.html>

Meissner, M. (2017) China's Social Credit System, <https://www.merics.org/sites/default/files/2017-09/China>

Sandhu, N. (2018). The Future of AI: It's is Not About A Computer That Thinks Like A Human. <http://blogs.blackberry.com/2018/11/the-future-of-ai-its-not-about-a-computer-that-thinks-like-a-human/>

Thomson, E. (2017). A Deontological Ethics Analysis of China's "Social Credit System". <https://earlsblog.me/2017/11/08/a-deontological-ethics-analysis-of-chinas-social-credit-system-part-1-the-background/>