

Homo semper maior: člověk v době AI

Tomáš Mančal

Příchod umělé inteligence (AI) není jako příchod jiných technologií. Stejně jako ostatní nové nástroje zvyšuje AI lidský potenciál, na rozdíl od nich však současně s člověkem soupeří v oblasti, kterou jsme si dlouho uzurpovali jen pro sebe – v myšlení. Setkání s AI nás nutí nově vymezit to, co je výhradně lidské mnohem skromnějším způsobem, který však může být zároveň inkluzivnější k nám samým a umožňuje znovu objevit zdroje naší výjimečnosti.

V předcházejícím čísle *Universa* jsme otevřeli téma dopadu umělé inteligence na vztah člověka k vědění a na to, co to znamená poznávat.[1] Nastínili jsme, že s umělou inteligencí (potenciálně schopnější než lidskou) možná přichází nutnost znovu promyslet podstatu racionality. Původním záměrem příspěvku ale bylo dojít v úvahách až k tomu, co znamená příchod umělé inteligence pro samotného člověka, pro jeho místo na zemi, zkrátka pro jeho lidství. V první části jsme alespoň dospěli k závěru, že se člověk nemůže dál stavět do role poloboha nadřazeného všemu světu z titulu své jedinečné schopnosti myslet. Máme dnes totiž mnoho důvodů domnívat se, že myšlení (alespoň ve smyslu toho slova zúženém na měřitelné intelektuální výkony) není zdaleka tak jedinečnou schopností, jak jsme se ještě donedávna domnívali. V tomto příspěvku se pak pokusíme rozvinout předchozí diskusi směrem k otázce, co vypovídá skutečnost umělé inteligence o nás samotných a jak může změnit vidění našeho místa ve světě.

AI jako principiální model myšlení

Poznávání, kterému jsme se věnovali v předchozím zamyšlení, člověka vnitřně formuje. Umělá inteligence tedy svým zásahem do charakteru našeho poznávání nutně také proměňuje způsob, jakým člověk definuje a interpretuje sám sebe. Způsob, kterým s AI interagujeme, je často velice lidský – můžeme si s ní psát a hovořit. V takové interakci jí možná automaticky přisuzujeme lidské vlastnosti a cítíme, že se nám v mnohém podobá. Ptáme se, nakolik je podobnost jen zdánlivá a nakolik může AI poskytnout platný model pro diskusi o lidství a lidské identitě.

Z hlediska hardwaru je zřejmé, že se nejedná o nápodobu lidského mozku. Naše počítače jsou však *univerzální*, a mohly

by tedy s pomocí fyziologických detailů simulovat procesy probíhající v mozku. Odlišný hardware tedy není tím zásadním ani nepřekonatelným. Dnešní AI ovšem ani zdaleka fyzikální a chemické procesy v mozku nesimuluje. Místo toho používá matematické operace na principiálně jednoduché, avšak z hlediska možností a flexibility obrovské vnitřní reprezentaci dat. Jazykovým jednotkám vstupu (tokenům) přiřazuje AI body ve svém reprezentačním prostoru a s nimi jako s geometrickými entitami pak provádí relativně jednoduché matematické operace, jejichž parametry vznikly trénováním. Tyto operace popisují naučené vztahy mezi tokeny, jejich význam a způsob, jakým budou transformovány do výstupu. Fascinující je, že program, který sleduje v principu stejný základní mechanismus transformace vstupu do výstupu, jakou provádí moderní komerční AI, může být překvapivě krátký. Zjednodušený transformový model může mít třeba 100 řádků v programovacím jazyce Python.[2] Zásadní rozdíl mezi takovým programem na hraní a skutečným velkým jazykovým modelem (Large Language Model – LLM) je především ve velikosti reprezentace problému a kvalitě tréninku. Samotná délka programu obecně nic zásadního o složitosti problému neříká, ovšem v tomto konkrétním případě program skutečně odpovídá poměrně jednoduché sekvenci matematických operací (známých snad všem absolventům kurzů matematiky na technických vysokých školách) nad velkými daty. Jediné zásadní omezení na malých počítačích spočívá v přípustné velikosti dat a rychlosti jejich zpracování. Přesto si lze dnes stáhnout do výpočetně silnějšího notebooku vytrénovaný jazykový model a popovídat si s ním nebo mu zadat různé jednoduché úkoly. Takový malý LLM vám dovolí smysluplnou konverzaci v několika jazycích, pomůže s programováním v několika programovacích jazycích a případně za vás vyřeší slovní úlohu z matematiky. Model představovaný pouhými dvěma gigabyty místa na disku[3] by pravděpodobně

zvládl zkoušky z matematiky na osmiletá gymnázia včetně jejich zadání v češtině, byť češtinu jako svůj jazyk neuvádí.

Přestože takové modely nejsou příliš silné v sociálních aspektech komunikace, nelze jim upřít, že „myslí“ poměrně přesvědčivě. V roce 1950 publikoval Alan Turing test (nesoucí dnes jeho jméno) navržený k určení, zda stroj k tomu účelu sestrojený myslí.[4] Test probíhá formou hry, kdy se lidský aktér v písemné komunikaci snaží zjistit, kdo z jeho dvou společníků v diskuzi je stroj a kdo člověk. Problém tohoto testu je zřejmý. Ani pokud stroj úspěšně přesvědčí lidského účastníka o své lidskosti, neznamená to, že má nějaké vnitřní prožívání, nebo že vůbec myslí lidským způsobem. Tato jednoduchá a platná námitka ovšem naráží na principiální absenci nějakého lepšího testu. Turing nejspíše navrhl maximum toho, co o myšlení můžeme zjistit. Vždyť jakou máme jako jednotlivci informaci

o myšlení a vnitřním prožívání ostatních lidí? Kromě toho, že se můžeme navzájem vnímat více než jen písemně, co víme o vnitřním životě ostatních lidí? Jak víme, že naše protějšky pouze nemanipulují symboly (mezi které může patřit i celá škála mimoverbální komunikace) bez nějakého vnitřního prožívání? Nepochybně se můžeme pouze domnívat, že jejich prožitky jsou ekvivalentní těm našim a o jejich vnitřním životě pak můžeme dělat závěry pouze introspekci vlastních pocitů. Žádné objektivní stanovení aktu myšlení ve stroji, a dokonce ani v našem bližním, nelze podat. Přesto našim bližním automaticky přiznáváme schopnost myslet a prožívat a sami často soucítíme s jejich fyzickou bolestí. Jsme stvořeni tak, abychom napodobovali ostatní členy společnosti. Trénink našeho vlastního mozku se odehrává napodobováním ostatních za nevyřčeného předpokladu, že všichni jsme uvnitř v zásadě stejní.

V případě kapacit dnešních přenosných počítačů by malý model zcela určitě Turingovým testem neprošel, ale velké komerční modely již zřejmě kapacitu projít tímto testem v omezených experimentálních podmínkách mají. Některé experimenty již tvrdí úspěšné překonání Turingova testu moderními LLM.[5] I kdyby bylo ale možné tyto testy zpochybnit, není žádný důvod domnívat se, že správně trénovaný jazykový model se v blízké budoucnosti nenaučí člověka v takovém testu oklamat. Ve chvíli, kdy budou modely svou komplexitou a rozsahem srovnatelné s lidským mozkem, budeme mít teoreticky v rukou prostředky experimentovat s možnými modely lidské mysli.

Aniž bychom však takové modely měli, už samotný fakt, že zdání myšlení lze dosáhnout zřetěžením poměrně malého

počtu matematických operací, byť nad velkým množstvím dat, je pozoruhodný. Momentální stav umělé inteligence významně podporuje představu, že lidské myšlení vzniklo jako vedlejší produkt nutného procesu učení, kterého jsou schopny, ba ke kterému jsou nuceny všechny organismy. I ty nejjednodušší organismy se učí, ať už na úrovni jedince, nebo na vývojové úrovni jako druh. Dnešní AI má překvapivě jednoduchou strukturu a její vzrůstající dokonalosti v nápodobě myšlení se dosahuje do značné míry hlavně zvětšováním velikosti modelu, v jakési analogii s růstem počtu neuronů v mozku. Vývoj AI naznačuje, že složité myšlení může povstat v systému s jednoduchými pravidly jen růstem jeho velikosti. Dnešní velikost lidského mozku musela být pro lovce-sběrače evoluční výhodou již před stovkami tisíc let. Lidský jedinec se po tisíciletí učil maximum toho, co mu byla schopna předat jeho kultura,

a teprve v posledních tisících let se k tomu přidaly abstraktní disciplíny jako literatura a matematika, které se prehistorický člověk neměl kde naučit. Tyto disciplíny vznikly jen pomalým sbíráním znalostí na úrovni větších lidských celků. Často se pozastavujeme nad tím, k čemu byla prehistorickému člověku schopnost myslet abstraktním způsobem, k čemu mu byla principiální schopnost odvodit Gödelův teorém nebo pochopit kvantovou mechaniku. Velmi plauzibilní odpovědí, kterou nám téměř experimentálně poskytuje AI, je, že mu tyto schopnosti klidně mohly být úplně k ničemu. Nikdy totiž samy o sobě nehrály v evoluci žádnou roli. Tyto schopnosti jsou emergentním projevem složitosti lidského mozku. Složitosti, která je sama o sobě užitečná i pro lovce-sběrače. Abstraktní myšlení tedy zdaleka nemusí být takový ojedinělý jev, jak jsme si doposud mysleli, a s příchodem AI tím oslabuje i výjimečnost a záhadnost lidské mysli samotné.

Kdo je člověk?

To nás znovu přivádí k aktu odsunutí člověka z jeho božského piedestalu, jak jsme se o něm zmiňovali v úvodu a v minulém čísle *Universa*. Zdá se, že se tu jedná o další z mnoha koperníkovských obrátů, kterými si lidstvo prochází. O další ztrátu pevné půdy pod nohama. Domnívali jsme se, že myšlení nám zajišťuje výjimečné místo ve vesmíru, tato výjimečnost však zmizela s příchodem AI. Země přestala být středem vesmíru. Přitom jsme podle této výjimečnosti hierarchizovali svoji společnost. Ti nejlepší z nás, kteří podávají nejlepší výkony ať už jde o business nebo vědu, nebo dokonce jen schopnost



Tomáš Mančal

Momentální stav umělé inteligence významně podporuje představu, že lidské myšlení vzniklo jako vedlejší produkt nutného procesu učení, kterého jsou schopny, ba ke kterému jsou nuceny všechny organismy.

→ manipulovat ostatními, se těší v naší společnosti největší úctě. A i kdyby se snad netěšili přímo úctě, alespoň požívají výhod, na které ti s menšími schopnostmi dosáhnout nemohou. V dnešní společnosti je stále obtížnější přicházet s argumenty, proč bohatství nabyté ekonomickým růstem zpřístupňovat těm, kteří si ho podle zúžených pravidel našich meritokratických společností nezasluhují. Je zcela běžné označovat ty, kteří nezapadají dobře do naší představy řádu světa, za hlupáky; první reakce na něčí nečekané chování nás může vést ke zvolání: „On se snad zbláznil!“ Dalo by se nalézt mnoho příkladů z minulosti a bohužel i ze současnosti, kdy celé skupiny lidí byly na základě národnosti, rasy, či prostě jiného názoru označovány za méněcenné, a toto označení šlo a jde často ruku v ruce s představou nižší mentální výkonnosti. Teď nám ovšem hrozí, že ve srovnání se schopnostmi umělé inteligence budeme hlupáky (zcela objektivně) všichni. Výjimečnost člověka jako myslícího tvora je tatam.

Máme snad tedy podle naší dosavadní praxe přisuzování větší hodnoty těm schopnějším přikládat nejvyšší hodnotu umělé inteligenci? Je snad pravda, že jsou to mentální schopnosti, které zakládají lidský nárok na úctu a důstojnost? V žádné společnosti, která ještě nepropadla úplnému sobectví, by ve skutečnosti takové hodnotové měřítko neuspělo. Pokud je naše nová situace skutečným koperníkovským obratem, pak lidská mentální výkonnost není měřítkem důstojnosti a lidskosti. Země sice přestala být středem vesmíru, ale na obloze se nic nezměnilo. Člověk se mýlil, když se domníval, že to je jeho schopnost myšlení, která mu dává důstojnost a která ho opravňuje považovat se za výjimečného. Člověk se ale nestavěl, kam nepatřil, pouze se tam stavěl z nesprávného důvodu a nový pohled na jeho situaci mu umožní nalézt pravdivější podobu sebe sama a dovolí dát jeho výjimečnosti skromnější a pravdivější význam.

Pro křesťana nemůže být tento náález žádným velkým překvapením. Ježíš ve svém horském kázání neslibil nebeské království chytrým. Nad významem spojení „chudí v duchu“, [6] kterým bylo v horském kázání přislíbeno nebeské království, se jistě můžeme dlouze pozastavovat. Je nejspíš správné přeložit οἱ πτωχοὶ τῷ πνεύματι spíše jako „ti, kdo mají duši chudého“, jak to dělá Jeruzalémská bible. Evangelium neodkazuje na intelektuální výkonnost, ale na duchovní nastavení, na nespolehání se na sebe sama, na odkázanost na Boha, zkrátka na nastavení toho, kdo nic nemá. Ale kouzlo významových posunů a vývoje jazyka nám dovoluje vzít vážně moderní, spíše kolokviální význam spojení „chudý duchem“, které nalézáme v Bibli kralické.

Tomáš Mančal

Pro budoucího člověka
a jeho schopnost vypořádat se
s nástrahami vlastního pokroku
bude zcela klíčové
jeho sebepochopení.

Před naším vlastním výtvořem, umělou inteligencí, můžeme klidně jednou být těmi chudými, nedostatečnými duchem, tedy schopností myslet. Chudí tou schopností, kterou jsme v jednu chvíli považovali za naši výjimečnou vlastnost. V naší dnešní společnosti, kde se materiální chudoba tak silně váže na dosažené vzdělání a dosažené vzdělání silně koreluje s materiálním zázemím rodin, [7] najednou nabývá pejorativně chápané kra-

lické spojení „chudý duchem“ odpovídajícího významu. Vždyť kdo je dnes chudý, nebude svým dětem schopen dopřát stejné vzdělání jako jeho ekonomicky úspěšnější spoluobčan. Nedostatek vzdělání vede zpět k chudobě a začarovanému kruhu, který plodí další sociální nerovnost. Nejde o lineární příčinnost, ale o kruh, který je obtížné prolomit bez podané ruky. Ruky,

kterou se společnost založená na sebeklamu důležitosti vlastní mentální výkonnosti a zásluh často zdráhá podat. Podle Ježíše bude ale právě těmto lidem patřit nebeské království.

Bůh, člověk, stroj

V naší vynalézavosti jsme si sestrojili nový velmi účinný nástroj, který nám slibuje ulehčit si život. Není to poprvé, co si ulehčujeme nikoliv od fyzické, ale intelektuální práce. Ještě nikdy ale nebyl náš nástroj lidem tolik podobný a schopen tolika činností donedávna vyhrazených jenom jim. Opustit myšlení jako znak lidství a lidské výjimečnosti je jedna věc, vzdát se myšlení, jako potřeby rozvíjet lidství, jeho naplnění a sebezáchovu je věc druhá. Podobně jako dnes ohrožuje nedostatek pohybu zdraví populace, může i mít nedostatek mentální činnosti ve společnosti zvyklé přenechávat myšlení strojům značný dopad na zdraví populace. Pro budoucího člověka a jeho schopnost vypořádat se s nástrahami vlastního pokroku bude zcela klíčové jeho sebepochopení. Člověk, který se považuje za výjimečného, má pro tento svůj pocit tendenci odsouvat to, co není výjimečné, na druhé místo, nestavět to do středu své pozornosti. Pokud se cítíme být výjimeční pro svou schopnost myslet, mohou se na onom druhém místě ocitnout všichni, jejichž mentální schopnosti pokulhávají, ať už skutečně nebo zdánlivě. Zavržení mentální výkonnosti jako měřítka lidskosti naopak dovoluje zrovnoprávnit vztahy mezi příslušníky lidského rodu. Zároveň je však potřeba udržet člověka na vrcholu důstojnosti, aby ho tam nenahradilo něco jiného a získané zrovnoprávnění se nakonec neproměnilo v úpadek všech.

Je obtížné říci, zda sekulární humanismus může najít dostatečné motivace k udržení člověka zbaveného intelektuální výjimečnosti na zmiňovaném vrcholu, na místě, kde by se mu

mělo dostávat největší pozornosti a důstojnosti. V čísle 4/2024 *Universa*[8] jsem načrtl jistou trajektorii vzestupu a pádu náboženství a vědy (či vědeckosti) jako schémat určujících a garantujících pravdu pro celé lidské společenství. Tam jsem vyjádřil názor, že po ústupu vědy už nebude možné opírat se o nic jiného než o člověka samotného, který už nemůže na nikoho (na Boha) a na nic (na vědu) odložit zodpovědnost za sebe a za naši planetu. Zdá se však, že člověk v této situaci bude mít mocného pomocníka. Pomocníka povstálého z lůna vědy, ale takového, který se nám může natolik podobat a natolik nás může převyšovat, že by se nám mohl stát náhražkou Boha. Pokud se budeme domnívat, že s vyšší inteligencí je možné nalézt lepší morální pravidla, založená na nějakém užítu vyhodnotitelném jen vyšší inteligencí, než je lidská, můžeme snadno odsunout na onu druhou kolej všechny ty, kdo neodpovídají utilitárnímu ideálu dokonalého myšlení a výkonnosti. Pokud se nám ideálem stane dokonalá umělá inteligence, nebudeme nikdy vzdáleni tomu, abychom se považovali pouze za (nedokonalé) myslící stroje. To, co si myslíme, že jsme nebo naopak nejsme, má obrovský vliv na to, jak chceme, aby s námi ostatní zacházeli.

Křesťané, alespoň pokud reflektují svoje náboženství a jeho kořeny, znají před zbožštěním AI mocná varování. Tím, že mohou lidskou důstojnost založit na faktu lidské stvořenosti Bohem, na jeho lásce k lidem a oběti jeho Syna, není pro ně ztráta výjimečnosti lidského myšlení vlastně žádnou podstatnou ztrátou. Ani pocit, že lidská mysl je nedokonalá, člověk omezený a že nad ním existuje něco, co ho značně převyšuje, není pro křesťany žádnou novinkou. Zároveň víme, že nic než Bůh nesmí hrát roli toho, kdo je výše než člověk. Nic z tohoto světa si nesmíme učinit modlou, ani sebedokonalější inteligencí.

Pro schopnost člověka být nadále svobodný je naprosto klíčové, aby v hodnotové hierarchii vždy figuroval člověk výše než stroj. Pro křesťany je porozumění této hierarchii mnohem snazší, protože zároveň neznamená nekontrolované postavení člověka do role měřítka všeho. Nad člověkem vždy stojí Bůh. Filosofie a náboženství k vývoji AI nijak viditelně nepřispěly,

ovšem moudrost nabytá za tisíce let v kontemplaci lidského místa ve světě bude pro zvládnutí výzvy AI nenahraditelná.

Člověk vždy větší

Naše schopnost vytvořit stroj myslící lépe než my sami staví člověka do velmi zvláštní situace. Člověk se vystavuje jistě krizi identity. Představa, že to, co zdánlivě dělá člověka člověkem, může být simulováno strojem, dovoluje myslet vážně myšlenku, že žádný člověk v tom speciálním smyslu, ve kterém si představujeme být stvoření Bohem, ani neexistuje. K těmto myšlenkám však není nutně potřeba AI. Člověk byl takto ateismem charakterizován odjakživa. Stále ale na vrcholu díky své schopnosti myslet. Případ AI je ale mimořádně silný a odsunutí člověka na druhou kolej, jeho

ztotožnění se strojem, a ztráta jeho vrozené důstojnosti skutečně hrozí. Výsledky přírodovědy a nyní také praktické výsledky počítačových věd vždycky zmenšovaly prostor, kam jsme mohli Boha, ale také člověka, „schovat“. Je ale tímto zmenšením Bůh nebo člověk nějak skutečně menší? Tradiční křesťanská formule říká: *Deus semper maior* – Bůh je vždy větší. „Jestli tomu rozumíš, není to Bůh“, jak říká sv. Augustin. Zdá se, že abychom uviděli člověka stvořeného k Božímu obrazu a lidství, které bylo možno v Kristu sjednotit s božskou přirozeností, musíme se na něj dívat stejným způsobem.

Autor je teoretický fyzik, působí na Matematicko-fyzikální fakultě UK.

LITERATURA A POZNÁMKY:

- [1] Tomáš Mančal: *Universum* 1, 2026, s. 26.
- [2] Takový program ochotně nejen vygeneruje, ale i vysvětlí přímo sama AI.
- [3] Použit byl model *llama3.2*, ke stažení na <https://ollama.com/library/llama3.2>.
- [4] Alan Turing, Computing Machinery and Intelligence. *Mind* 59 (1950), č. 236 s. 433.
- [5] Cameron R. Jones a Benjamin K. Bergen, *Large Language Models Pass the Turing Test*, arXiv:2503.23674.
- [6] Mt 5,3, *Bible: Český ekumenický překlad*. Česká biblická společnost, 2008.
- [7] Daniel Prokop: *Slepé skvrny: O chudobě, vzdělávání, populismu a dalších výzvách české společnosti*. Host, 2020
- [8] Tomáš Mančal: *Universum* 4, 2024, s. 30.

křesťanství
A SVĚT

historie
A SOUČASNOST

literatura
A VÝTVARNÉ UMĚNÍ

komentáře
A ANALÝZY

recenze
A ROZHOVORY

PERSPEKTIVY

PŘÍLOHA KATOLICKÉHO TÝDENÍKU

www.katyd.cz