

„...umělá inteligence však dle mého takových schopností nedosahuje. Zatím.“

Zatím. To slovo tam každý připojíme – nedá nám to. Možná ho vyslovíme s jakousi nadějí v hlase, s dychtivým očekáváním toho, co neustále se zrychlující technologický vývoj přinese; možná nám v hlase zazní nejistota či obava. Bleskový vzestup umělé inteligence v posledních letech mnohé z nás trochu zaskočil. Jde o jedno z nejdiskutovanějších témat současnosti, a to nejen v oblasti technologie – a přesto se dosud mnohým z jeho aspektů nedostává tolik pozornosti, kolik by mělo. Než však zabředneme do dalekosáhlých úvah o neméně dalekosáhlých – byť prozatím z velké části nezmapovaných – důsledcích rychlého rozvoje umělé inteligence v posledních letech a než se dostaneme konečně i k samotné otázce, proč bychom AI mohli či nemohli využít při tvorbě eseje, zkusme si nejdříve stručně ujasnit celkový význam tohoto objevu pro lidstvo a co mu předcházelo.

Začnu jaksí hegelovskou myšlenkou. Jak už to tak s velkými milníky v lidském vývoji bývá, člověk žijící uprostřed určitého procesu často není schopen adekvátně posoudit jeho skutečný význam – vždyť kolik vynálezů, myšlenek, objevů bylo zprvu odsuzováno jako bezvýznamnosti, a dnes na nich jako civilizace stavíme a považujeme je za nedílnou součást toho, co jsme dokázali? Nebo naopak: kolik okamžiků bylo v lidských dějinách oslavováno, prohlašováno za naprosto průlomové, a dnes se o nich dočítáme jen na stránkách učebnic dějepisu – pokud vůbec – a shledáváme je v lepším případě zamyšlení hodnými, většinou však jen úsměvnými? Jestli se z oné věci stane přelomový okamžik v lidské historii, či kuriozita v poznámce pod čarou, běžný člověk „v běhu dějin“ mnohdy nemá šanci poznat. Potřebuje odstup. Bez něj je slepý.

Existují však zvláštní případy, u kterých pro posouzení významnosti situace odstup nutný není. Už nyní třeba mnohým přijde nadmíru jasné, že příchod umělé inteligence se ke druhé kategorii událostí, tedy ke chvilkovým senzacím, jež budou po pár letech zapomenuty, téměř jistě nezařadí. Rozličné fantazie o lidmi stvořených entitách s vlastním vědomím a s potenciálem se proti lidem vzbouřit a nahradit je, nás provázejí odedávna. Od „mírnějšího“ příkladu Pygmalionovy Galatey, u které pravděpodobnost jakéhokoliv násilného činu proti svému tvůrci není závratná, přes pověsti o Golemovi až po slavnou Čapkovu hru R.U.R., prezentující vskutku odstrašující možnost vývoje robotů. Nyní se však možná nejdůležitější článek těchto představ, a sice schopnost námi stvořené technologie samostatně „myslet“, konečně začíná měnit ve skutečnost. Právě díky schopnosti myslet, či lépe řečeno přijímat informace a efektivně s nimi pracovat podle zažitých vzorců vytvořených lidmi, se objev umělé inteligence může zařadit mezi nejvýznamnější události lidské historie. Proto, že je zcela jiné povahy než cokoliv, s čím lidé do této chvíle přišli. Každý takový velký skok v technologickém vývoji ale také klade obrovské nároky na lidskou mentalitu a mnohdy vyžaduje zásadní změnu způsobu přemýšlení o světě.

Pocity mnoha lidí vyvolané příchodem umělé inteligence by se lehce daly přirovnat právě k prvotním pocitům již zmíněného sochaře Pygmaliona, zjistiťšího, že jeho milovaná socha, dosud chladná a

bledá, vytesaná z mramoru, se mu pod rukama začíná měnit v živoucí bytost. Tak je popisuje Ovidius ve svých Proměnách:

*„Žasne a nejistá radost jím prochvívá, klamu se bojí;
lásky však pln on zkouší opětně předmět své touhy:
tělo to jest — již tepají cévy, jichž prsty se dotkly!“¹*

Po tisíciletí jsme vytrvale z mramoru tesali jednu sochu za druhou, tento výtvar se však vymyká všemu, na co jsme zvyklí. Naše vlastní socha nám teď pohlédla do očí, námi prochvívá nejistá radost a nevíme, jak s tím naložit. Je to jen na nás.

Nyní by se patřil obligátní moralistní apel na to, jak ohleduplně a obezřetně si musíme jako generace i lidstvo v tuto klíčovou chvíli počínat a jak musíme mít na paměti všechny výhody, ale i nevýhody a potenciální hrozby této technologie. Jak je naší volbou, jestli se umělá inteligence stane Galateou, či Golemem. Význam tohoto apelu bych sice nepodceňoval, sem ho však nezařadím. I proto, že už jsme ho všichni slyšeli tolikrát a je to zkrátka fakt; nezbyvá už mnoho co doplnit. Snad jen ještě krátce okomentovat následující, zprvu trochu zarážející skutečnost, úzce související s výše zmíněnými tématy.

Když se jednou za čas objeví výrazný fenomén s rozsáhlým celospolečenským dopadem, dokáže si průměrný člen dnešní globalizované společnosti na toto téma s ohromující rychlostí vytvořit svůj vlastní názor – občas velmi vyhraněný. Může se zdát, že jde o jakousi nepsanou občanskou povinnost (další otázkou samozřejmě je, zda onen člen společnosti informace pro své stanovisko čerpá ze správných zdrojů – na tom teď ale nesejde). Nyní, vprostřed debat ohledně AI, člověk však až nezvykle často naráží na lidi, kteří se dosud neodvážili si vlastní vyhraněný názor vytvořit – další důkaz toho, že jde pro společnost o velkou neznámou. Pořád moc nevíme, co si o tom máme myslet, a to dává nově příchozí technologii čas pro zapuštění kořenů. O možných etických důsledcích této situace však později.

Máme-li se dále zamyslet nad rozličnými dopady umělé inteligence na společnost, v našem případě zvláště na oblast tvůrčího psaní, jednou z nejočividnějších otázek je, nakolik umělá inteligence změní pojem autorství. Hermann Hesse ve své Hře se skleněnými perlami popisuje dystopickou společnost se značně deformovanými kulturními a duchovními hodnotami a před možností takového vývoje varuje. V knize je mimo jiné právě i postupné vymizení konceptu autorství při tvorbě umění jedním ze symptomů úpadku lidské kultury – toto lze velmi dobře pozorovat i na naší historii, budeme-li období středověku považovat z širšího hlediska za dobu všeobecného kulturního úpadku, což bez větších výčitek můžeme. V té době se téměř úplně ztratil význam identity autora; podívejme se na jakýkoli středověký spis a jméno autora často vůbec nenajdeme. Podobná situace může nastat ve chvíli, kdy se umělá inteligence stane naprosto běžným nástrojem pro tvorbu textů. V tom okamžiku se autor stane jen jakýmsi zprostředkovatelem mezi čtenářem a bezednou záplavou informací tekoucí z digitální sféry. Všimněme si, že to stejné lze říci i o chvíli, kdy se stal všeobecně přístupným internet

– nyní vskutku jde o to samé, ale přeci jen na trochu jiné úrovni, neboť umělá inteligence nám internetové zdroje „předžvýká“ a servíruje přímo.

Umělá inteligence však také přímo zpětně ovlivňuje prostředí, ze kterého vzešla; jelikož je každý jazykový model trénován na limitované množině vstupních dat, není nikdy schopen pokrýt jazyk v celé jeho rozmanitosti, a tím pádem se teoreticky nikdy ani nemůže přiblížit přirozenosti lidského projevu. Z téhož důvodu také začne s rostoucím počtem uživatelů čím dál více formovat podobu jazyka samotného. Jazykový model bude mnohem častěji používat slova, fráze a větné struktury, jejichž výskyt je ve vstupní množině dat frekventovanější. Texty těmito modely vygenerované začnou přirozeně používat i lidé a po čase se na nich začnou trénovat nové jazykové modely. A už tu máme Urobora požírajícího vlastní ocas – čili umělou inteligenci trénující se na textech, jež sama vytvořila. Pomalu, ale jistě je tak schopna omezovat rozmanitost slovní zásoby nejen textů, které generuje, ale posléze i samotného jazyka, ve kterém je generuje (samozřejmě za předpokladu, že do té doby lidé nepřijdou s mechanismem, jenž by byl schopen tento proces eliminovat). Tím pádem se umělá inteligence sama stává činitelem onoho zmiňovaného kulturního úpadku. Toho, jež Hesse přirovnává k *„trouchnivění hluboko v kmeni stromu lidského duchovního bohatství, dlouho zůstávši nepovšimnuté, jež postupně se rozrůstá a brání stromu ve spokojeném růstu, až nakonec přeruší kanálky s životadárnou mizou a strom, po staletí rostoucí, pomalou smrtí zevnitř skoná – ačkoliv navenek se ještě značnou chvíli bude zdát zdravý a neporušený“*.² Inu doufáme, že tak zlé to nebude.

Ted' něco k etické stránce používání umělé inteligence. Jak jsem již naznačil, obdobné fenomény mnohdy bezprostředně po svém příchodu vyvolají ve společnosti nějakou chvíli trvající období, kdy si ještě společnost nestihla vytvořit všeobecně uznávaný soubor morálních stanovisek onoho tématu se týkajících – vzniká tedy chvíle, v níž některé nezbytné principy chybějí, čili jakési „morální vakuum“. Tento koncept ukážu na následující metafoře: představme si společnost jako vrátného na noční směně. Morální principy společnosti jsou zde představovány vůlí vrátného nevpustit do budovy nikoho, kdo by mohl napáchat škodu. Ono zmiňované „morální vakuum“ je zde moment, kdy mysl vrátného zláká slastné spočinutí ve snu, ale dveře, jež předtím nezamkl, zůstanou na štěrbinku pootevřeny. V onom momentě může do budovy vniknout noční vánek a spolu s ním i kdokoli, kdo by zrovna měl cestu kolem – avšak jen do chvíle, než se průvan obrátí a dveře se s prasknutím, ještě několik mučivě dlouhých vteřin se rozléhajícím po spící budově, zavřou. Toto prasknutí, samozřejmě způsobivší poněkud násilné probuzení nebohého vrátného, představuje okamžik, kdy se společnost konečně vzpamatuje a vytvoří si na problematiku více či méně jednoznačný názor. Jinými slovy „morální dohled“ společnosti se v tomto bodě vrátí, stejně jako snaha zmateného, provinilého a rozespalého vrátného co nejdříve dveře zabezpečit.

Toto „morální vakuum“, tedy okamžik, kdy dvířka zůstanou otevřená a bez dohledu, lze přirozeně vnímat jak pozitivně, tak negativně. Jako dobu, kdy se zkostnatělé zvyklosti společnosti nestačí sdostatek rychle přizpůsobit pokroku, a kdy se tedy otevřou zcela nové možnosti pro reformu a takřkajíc „nový život“ – v naší metafoře by šlo o svěží jarní noční vánek vnikající do zatuchlé budovy. Nebo naopak jako chvíli, kdy léty ověřené etické principy společnosti na okamžik jaksi nejsou dostačující pro právě vzniklou situaci, a tak se pomyslná dvířka v mezidobí, než se společnost vzpamatuje a vytvoří si nové stanovisko, otevírají potenciálně škodlivým silám, které jsou v krajním případě s to napáchat nevratné škody a stát se spouštěči úpadku. Tento případ by pak odpovídal třeba mrazivému podzimnímu větru, mařícímu příjemné klima dobře vytopené budovy.

Zde svůj tok myšlenek přeruším, neboť další stránka nesouvislého blekotání o abstraktních pojmech už by nikoho nebavila. Podívejme se teď raději na otázku, k níž celou dobu směřuji: může za mě esej napsat umělá inteligence? Při odpovědi na tuto otázku se vedle všech faktorů, které jsem již uvedl, z praktického hlediska zdají býti rozhodujícími tři hlavní okolnosti:

Zaprvé, používáme-li při psaní umělou inteligenci jako autora či jako korektora. Pokud AI napsala text celý a náš jediný autorský vklad bude napsání promptu do zadávacího řádku, pak se s čistým svědomím zkrátka za autory v tradičním slova považovat nemůžeme, ať naše obhajoba bude sebezpřesvědčivější. Až se však umělá inteligence naučí spolehlivě ovládat pravidla české gramatiky a pravopisu, stane se její široké využití na poli korektury nevyhnutelným – podívejme se do anglicky mluvících zemích, kde k tomuto již částečně došlo. A využití umělé inteligence ke kontrole a korekci formální stránky textu dle mého názoru není třeba nijak morálně legitimizovat. Pro potřebu dalších úvah tedy pod slovy „využití AI k tvorbě textu“ uvažujme to, že necháme celý text nebo jeho většinu napsat umělou inteligenci a náš tvůrčí vklad bude minimální.

Zadruhé, za jakým účelem onu esej píšeme. Čili jestli je pro nás prostředkem sebevyjádření, sdělení vlastního názoru a utřídění myšlenek, nebo jen dalším z řady školních úkolů – či vzletněji řečeno akademických povinností. Jestli onen text píšeme z vlastní vůle, nebo zkrátka proto, že musíme, a tím pádem chceme mít proces psaní co nejrychleji odbytý. V prvním případě je závěr poměrně jednoznačný – proč bychom nechávali text napsat umělou inteligenci, když tím svým způsobem popíráme samu podstatu textu, který píšeme? Pokud je ale pro nás psaní eseje pouhou povinností – třeba právě v podobě školního úkolu – a my usoudíme, že pro nás nebude nijak přínosné, když se nad tématem eseje opravdu zamyslíme a napíšeme ji sami, pak je dle mého názoru využití AI při psaní takové eseje (samozřejmě z hlediska, jež se povznese nad zásady a pravidla školního prostředí) naprosto ospravedlnitelné.

A konečně zatřetí, a sice jestli sami následně uvedeme, zda jsme umělou inteligenci využili. Asi se shodneme na tom, že postup, kdy necháme text napsat umělou inteligenci a poté ho vydáváme za svůj vlastní, je nejen stěží morálně obhajitelný, ale také v reálném životě stěží použitelný, neboť text vygenerovaný AI vykazuje jisté charakteristické znaky, díky nimž ho lze nepříliš obtížně identifikovat. V praxi umělá inteligence tímto Turingovým testem ve většině případů neprojde. Zatím.

Co nám z uvedených úvah vyplývá? Umělá inteligence, jsa sama výsledkem pokroku, se může stát buď činitelem ještě rapidnějšího a masivnějšího pokroku – bude-li dobře využita –, nebo naopak činitelem neméně masivního úpadku. Tato formulace je však natolik vágní, že se neobejde bez několika upřesnění: kupříkladu co přesně znamená dobře využita? Dalo by se říci, že všechny vynálezy, s nimiž do této chvíle lidstvo přišlo, lze popsat jako věci, které člověku ulehčují či zefektivňují práci tak, aby měl čas věnovat se přínosnějším a smysluplnějším činnostem. Člověk pak neztrácí čas otrockou prací a nemusí se tak plýtvat potenciálem lidské mysli, který by mohl být využit nepoměrně přínosnějším způsobem. Tento princip lze samozřejmě sledovat na mnoha úrovních důležitosti, jeho podstata ale zůstává zachována. I v souvislosti s AI lze teď už tento názor považovat za všeobecně přijímané stanovisko společnosti. Podle mnohých by tedy ideální postup, jak zacházet s umělou inteligencí, byl prostě nepouštět AI k oborům lidské činnosti, v jejichž samotném jádře stojí

lidská přirozenost – tedy k umění, literatuře, poezii, hudbě, filosofii, některým aspektům vědy –, a nechat ji pracovat na věcech, které nás do této chvíle jen omezovaly. Které dosud vyžadovaly příliš tvůrčí invence na to, aby je mohl dělat stroj, ale pro lidi byly příliš jednotvárné a nevytěžující, takže u nich člověk ztrácel čas, jenž by býval mohl být využit mnohem smysluplněji.

Bylo by však omezené tvrdit, že je možné, abychom umělé inteligenci ze dne na den odřekli přístup k oborům, jako je umění nebo hudba, a přestali ji v těchto oborech úplně používat. Umělá inteligence již do těchto oborů zabrousila a nechala za sebou nesmazatelnou stopu. V tuto chvíli je dle mého mnohem lepší se zaměřit na to, jak konkrétně můžeme AI využívat tak, jak jsem naznačil; tak, aby z nás sundala břímě zbytečné práce navíc, která jen mrhá lidským potenciálem. Využijeme-li ji takto, umělá inteligence může, v kombinaci s relativně nedávno vynalezeným internetem, nakonec mít srovnatelný dopad na kulturní vývoj lidstva jako třeba vynález knihtisku – když už opisovači nemuseli trávit dlouhé hodiny ručním šířením textů, rapidně to urychlilo šíření informací a přispělo k všeobecné obrodě lidské kultury v polovině minulého tisíciletí. Podobný průlom byl vynález počítačů a kalkulaček; zařízení, která dokážou s velkou přesností a v krátkém čase provádět komplexní výpočty. Tyto vynálezy nás vysvobodily z otrocké práce a daly tak prostor pokroku zcela jiných rozměrů. Tudíž chceme-li sledovat jakési všeobecné zájmy lidstva, i naše počínání s umělou inteligencí by se v následujících letech mělo ubírat tímto směrem.

Čeká nás tedy v následujících desetiletích novodobá renesance? To si odhadnout netroufám – koneckonců člověk v běhu dějin není s to plně poznat souvislosti událostí, jichž je sám součástí. Každá událost je jen výsledkem nepřehledné změti různých souvislostí, příčinných vztahů, důvodů a důsledků, a sama se následně součástí této změti stává. Jak by řekl Hegel: „...*dějiny se zprvu jeví jako poslušnost náhodných událostí, v níž stojí každý fakt izolovaně sám o sobě, a je možno poukázat jen na souvislost časové následnosti. My však poznáváme v nich zákonitý vztah, jímž nabývají jednotlivé události svého zvláštního postavení, svého vztahu k určitému cíli nebo účelu, a tím i svého významu. Historický význam záleží toliko ve vztahu k obecnému. Poznat obecné znamená pochopit význam.*“³

Prohlašuji, že jsem při vzniku tohoto příspěvku nepoužil umělou inteligenci.

Citace:

¹ Naso, Publius Ovidius. *Proměny*. Městská knihovna v Praze 2022.

² Hesse, Hermann. (1943). *Hra se skleněnými perlami*. Argo 2012.

³ Hegel, Georg Wilhelm Friedrich. (1837) *Dějiny filosofie*. NČAV 1961.