

Budoucnost naší budoucnosti

Jan Romportl vystudoval kybernetiku a umělou inteligenci na Fakultě aplikovaných věd ZČU a filozofii vědy na Fakultě filozofické. Vede centrum pro umělou inteligenci společnosti O2 Czech Republic a na ZČU se vrací jako host s přednáškami na téma **Budoucnost**.

Začátkem letošního roku – shodou okolností zrovna v době nástupu první pandemické vlny covid-19 – vyšla dlouho očekávaná kniha od filozofa Tobyho Orda z Future of Humanity Institute (FHI) na Oxfordské univerzitě. Kniha se jmenuje *The Precipice*, tedy *Propast*. Toby Ord není žádný akademický „hej-počkej“. Je to brilantní myslitel vycházející z prostředí morální filozofie, ale s obrovským vhladem do matematiky, informatiky a přírodních věd. Mimo jiné stál u zrodu globálních hnutí *Effective Altruism* a *Giving What We Can*. Cílem jeho bádání není nic menšího než hledat odpověď na otázku, jak zajistit co největší dobro a co nejdlejší budoucnost pro lidstvo.

Troufám si říci, že kdybychom si tento cíl výslovně stanovili na některém ze „standardních“ českých univerzitních pracovišť, vysloužili bychom si od jedné části mainstreamové akademické obce úsměšky a tukaní si na čelo, od druhé části pak úsměšky a osočení z nabubřelosti. A přitom právě tento cíl by měl být tím nejdůležitějším, ke kterému by každá společnost (a tedy i univerzita jako jedna z klíčových společenských institucí) měla mířit. Nikdo by se tedy neměl stydět za to, otevřeně se k němu přihlásit a směřovat k němu své osobní i institucionální kroky.

Kniha Tobyho Orda se zabývá tzv. existenčními riziky. V současnosti je asi nejucelenějším pojednáním o této problematice. Řekl bych, že Tobyho cílem bylo napsat takovou knihu, která pro oblast zkoumání a minimalizace existenčních rizik bude tím, čím pro oblast umělé inteligence a její bezpečnosti byla již před šesti lety kniha *Superintelligence* od Nicka Bostroma (taktéž z FHI).

Existenční rizika jsou takovou vyšší a děsivější úrovní globálních katastrofických rizik. Dalo by se říci, že globální katastrofická rizika má lidstvo ještě šanci „rozchodit“ (byť třeba za cenu enormních ztrát lidských či materiálních, které nás mohou vrátit klidně o několik staletí zpět), ale nezvládnutá existenční rizika lidstvo už „nerozchodí“. Přesněji: existenční rizika jsou takové události, které trvale zničí dlouhodobý potenciál lidstva. Ve většině případů jde o úplné fyzické vyhlazení lidstva, ale může jít i o scénáře jako dystopická superstabilní totalitární společenská zřízení, z jejichž Nashových ekvilibrií se již nelze vytrhnout (něco jako Orwellův 1984, ale ještě horší), nebo třeba vyčerpání veškerých zdrojů energie a technologický návrat na úroveň neolitu.

Toby Ord velmi důsledně mapuje různé druhy rizik a snaží se o pečlivé Fermiho odhady rozsahu jejich



Jan Romportl

možných dopadů i toho, s jakou pravděpodobností mohou nastat. Ukazuje, že pravděpodobnost antropogenní (tedy člověkem způsobené) existenční katastrofy o mnoho řádů převyšuje rizika ne-antropogenní. A dochází k velmi alarmujícímu závěru: šance, že během příštích 100 let dojde k antropogenní existenční katastrofě, je podle něj 1:6. Doslova tedy ruská ruleta o budoucnost lidstva.

A zde se právě dostáváme k pointě celé knihy: byť je prognóza děsivá, jde o antropogenní rizika. Tedy ta, která si vytváříme sami. A proto je zároveň plně v naší moci jim zabránit. Nejsme zde jako pouhé kachny na odstřel – budoucnost naší budoucnosti je totiž plně v našich rukách. Ale bude to extrémně složité. Jsme ve zcela unikátní situaci v dějinách života na Zemi: jako první živočišný druh máme schopnost aktivně zničit sami sebe a patrně i většinu ostatního života, ale zároveň jako první živočišný druh máme teoretickou schopnost tomuto zabránit a naopak během dalších staletí a tisíciletí rozšířit inteligentní život za hranice Země.

Dlouhodobý potenciál lidstva je nepředstavitelně větší než ten, který jsme rozvinuli doposud. Stojíme na samém jeho počátku, avšak rovnou

na bezprecedentním historickém milníku – doslova nad *Propastí*: naše globální akce během několika příštích desetiletí rozhodnou o výsledku prvního kola ruské rulety. Buď první část přechodu nad *Propastí* zvládneme a otevřeme si významnější díl budoucnosti, nebo klopýtáme, do *Propasti* spadneme a ztratíme víc, než co kdy lidstvo kumulativně ztratilo doposud.

Kdo jiný než univerzita by měl připravit novou generaci lidí, aby byli schopni se takového úkolu zhostit?

Je na to ale univerzita připravena? Myslím si, že není. Ani ta naše, ani žádná jiná česká. Proč? Protože by si univerzita jako svůj hlavní cíl musela budoucnost lidstva beze studu přiznat. Protože by univerzita musela přestat věst žabomyši vnitřní války a tribunální spory mezi různými týmy, katedrami či ústavami. Protože by se musela mnohem více zaměřit na své hlavní poslání – na systematické kvalitní a univerzální vzdělávání směrem k idejím racionality (jak se o ní dočtete třeba na lesswrong.com) a otevřené společnosti, namísto toho, aby budovala skupinky fuchidiotů slepých ke všemu mimo svůj vymezený obor. Protože by univerzita musela umět podpořit upřímně a bez závisti svůj mezinárodně excelentní výzkum, ale zároveň najít odvahu zastavit řadu výzkumných snah, které zápasí v úrovni okresního přeboru na hranici smysluplnosti i finanční rentability, prostě jen tak ze zvyku. *Non multa sed multum*. Vždyť je přece spousta jiných zajímavějších a užitečnějších činností než unavené grantobijectví a hyperoptimalizace publikačních zářezů do RIVu. Třeba kvalitní vzdělávání a záchrana lidstva.

Myslíte, že to univerzita dokáže? Podobně jako u těch antropogenních rizik – je to v našich rukou... ●

Nashova ekvilibria – Americký matematik John Forbes Nash se zabýval teorií her. Nashova rovnováha (*Nash equilibrium*) je situace, při níž nemůže žádný z hráčů jednostrannou změnou zvolené strategie vylepšit svoji situaci.

Fermiho odhady – Italský fyzik Enrico Fermi byl známý překvapivě velmi přesnými odhady na základě malého množství údajů.