

Posudek oponenta bakalářské práce

Název práce: Umělá inteligence a Superinteligence

Autor práce: Michal Krňák

Vedoucí práce: Richard Zika, Ph.D.

FHS UK 2023

Posuzovaná práce je věnována jednomu z nejvíce konjunkturních témat současnosti, sice kritickému zkoumání perspektiv vývoje umělé inteligence. Perspektiv v tom smyslu, že se práce zabývá jak optimistickým (zde především pod hlavičkou transhumanismu a v podání Borstromově), tak pesimistickým pohledem na to, co lidstvu umělá inteligence může přinést. V tomto smyslu je hlavním sdělením práce základní vymezení pojmů, koncepcí a teoretických rámců, které tematizují vývoj umělé inteligence od počátků počítačové vědy a přes strojové učení, neuronové sítě, atp. až k Superinteligenci.

Práce je solidní a neshledávám v ní něco vyloženě významně problematičtějšího, byť – a to zvláště z pozice oponenta – nacházím řadu dílčích problematičtějších bodů. Obvykle se jedná o menší významové posuny, jež však po mém soudu vedou k desinterpretacím některých děl:

S. 3: „Arendtová tak v podstatě hovoří v souladu s bájí o Prométheovi. Každý lidstvu prospěšný vynález a jeho praktické využívání je pak posouváním naší lidské přirozenosti.“ V jakém smyslu? O jakou přirozenost má jít v případě Arendt?

S. 14: Čím je Turingův test zastaralý? Samotné podání imitační hry neodpovídá Turingovu vlastnímu představení.

Dále k témuž: Známý spor Searla s Turingem je podán velmi jednostranně. Lze prokázat, že i zpětně poskytuje Turingův článek kritiku Searlova domnělého rozbití představy o (umělé) inteligenci, kterou zastává Turing. Proto ani neplatí, že by se argumentem čínského pokoje rušila platnost Turingova určení inteligence.

S. 16: Searle mluví o sémantice (oproti syntaxi), nikoliv sémiotice.

Na trochu obecnější rovině se domnívám, že práce sleduje zároveň dvě poměrně rozporná stanoviska. Na jedné straně totiž interpretuje možnosti transhumanismu jako vylepšování člověka. A potud je umělá inteligence nástrojem či možností takového vylepšení.

Na straně druhé však řeší problém umělé inteligence jako otázku toho, co se stane, když bude umělá inteligence samostatným aktérem, resp. získá vědomí/svobodnou vůli atp. Z tohoto druhého hlediska je však vylepšování člověka irelevantní.

Nemohu si ani odpustit otázku, jak vlastně práce rozumí pojmu inteligence. Turingovo řešení je odmítnuto (i když si myslím, že jen proto, protože není adekvátně představeno a pochopeno), ale nejsem si vědom toho, že by práce poskytla lepší alternativu.

Práce je po formální stránce relativně v pořádku, moje jediná výtky, již si neodpustím, míří k anglicismům, zvláště pak otrockému překladu „myslící agenti“ (na mnoha místech), nebo „stanovení nějaké jednoduché optimy v jednání“ (s. 29).

Autorovi pak doporučuji přečíst si Lemova *Golema XIV*.

Práci považuji za velmi dobrou a doporučuji ji k obhajobě.

V Praze 10. září 2023,

Jakub Marek, Ph.D.