

POSUDEK DIPLOMOVÉ PRÁCE

Typ posudku: *oponent*

Autor/ka práce: Michaela Mendelová

Název práce: Humans, artificial intelligence and sentience

Vedoucí práce: Mgr. Filip Vostal, Ph.D.

Oponent/tka: Mgr. Martin Tremčinský, Ph.D.

Navržené hodnocení: **D-E**

Zdůvodnění Vašeho hodnocení práce (zejména výhrad a kritických připomínek) rozepište podle níže uvedených kritérií.

1. Je cíl práce (výzkumná otázka) jasně formulován a odpovídají mu závěry? Je práce vhodně strukturována?

Práce autorky si klade za cíl zodpovědět dvě hlavní výzkumné otázky: „whether US society is divided into groups relating opinions about the welfare of artificial intelligence and to describe what are these groups like“ (str. 3). Dále v textu jsou tyto otázky nazývány hypotézami:

„H1: Is US society divided into groups relating opinions about artificial intelligence?

H2: Is age, political preference, another sociodemographic, or experience with AI the dividing factor?“ (str. 9)

Není pak jasné, jak mohou být takto znějící hypotézy testovány. Zároveň takováto formulace vytváří v textu neurčitosti, jak např. na str. 8: „This further solidifies my hypothesis regarding demographic criteria,“ kde se čtenář musí domýšlet, jaké je znění této hypotézy.

Dále se v textu objevuje ještě otázka: „My question is whether the consumption of such media content (specifically the frequency) has any impact on the support of the well-being of sentient artificial entities“ (str. 7), která ale v textu není zodpovězena ani dále rozváděna.

2. Opírá se autor/ka o dostatek literatury relevantní tématu práce? Využívá i cizojazyčné texty?

Ano.

3. Jaká je kvalita použitých dat nebo zdrojů (včetně jejich výběru, sběru a popisu) a metod jejich analýzy?

Autora využívá sekundární analýzu dat vyprodukovaných amerických ve studii AIMS z roku 2021. Metoda analýzy LCA je podrobně rozepsaná a zdůvodněná.

4. Jaká je kvalita argumentů, o něž autor/ka opírá závěry, k nimž dospěl/a?

Autorka na základě analýzy identifikuje několik latentních tříd s odlišnými přístupy k právům umělé inteligence. Na základě těchto interpretací pak doporučuje strategie, jak komunikovat benefity etického přístupu k právům umělé inteligence. Z textu ale není jasné

- 1) komu jsou tato doporučení určena? (mediím, policy makerům, vývojářům AI?)
- 2) jaké jsou tyto benefity?

Autorka totiž jen v omezené míře reflektuje vlastní stanovisko ke zkoumané problematice, ale v interpretaci dat z tohoto stanoviska významně čerpá. Díky tomu pak vynáší nad respondenty výzkumu soudy, že nejsou dostatečně informováni o problematice umělé inteligence (str. 31), ale není jasné, co v tomto kontextu znamená být "dostatečně informován"? Zároveň o pár vět později tuto interpretaci autorka popře větou: „Nonetheless, maybe I am wrong, and they know everything they needed for their decisions.“ (str. 31-32) Opět, na základě čeho je možné usuzovat, že by vědění respondentů mělo být nedostatečné k vytvoření úsudku? Obzvláště když svůj úsudek vyjádřili v dotazníku, který práce analyzuje?

Celkově tak chybí přesvědčivé zdůvodnění, proč autorka považuje téma za společensky důležité, obzvláště když, jak sama uvádí, respondenti jej vesměs za důležité nepovažují (str. 32). Nechci zde tímto pochybovat o důležitosti samotného tématu práv umělé inteligence, ale poukázat, že text jej dostatečně nekomunikuje, ale spíše jej předpokládá jako všeobecně sdílenou morální pozici.

Jediného zdůvodnění se čtenáři dostane na str. 3, kde autorka uvádí: „The overall topic [předložené diplomové práce] is the public opinion on well-being of (sentient) artificial entities, therefore understanding the roles of these entities in society is important,“ kde tato formulace ovšem působí vyloženě nešťastně, protože vyznívá, jako že téma je důležité, protože si ho autorka zvolila a nikoliv že si jej zvolila, protože by bylo důležité.

5. Jsou v práci autorova/autorčina tvrzení a zjištění jasně odlišena od tvrzení a zjištění převzatých?

Ano, autorčina tvrzení jsou jasně oddělena od tvrzení převzatých.

6. Jaká je úroveň odkazového aparátu, jazyka a dalších formálních náležitostí?

Textu by prospěla důkladnější korekce. Chybí odsazení řádku na začátku odstavce, mezi některými odstavci jsou mezery, mezi jinými nikoliv. Na některých místech textu (hlavně v úvodu) se objevují věty, které působí neúplně (nejčastěji chybí větě přísudek).

7. Jiné přednosti a/nebo nedostatky, které neodpovídají výše uvedeným kritériím (jsou-li jaké).

8. Náměty na diskuzi při obhajobě práce.

Mohla by autorka zodpovědět některé otázky, které už zazněly v tomto posudku? Jmenovitě komu jsou určena doporučení v závěru textu a z jaké pozice autorka vychází? Proč je podle ní téma důležité a je nutné jej společensky moderovat?

9. Prohlašuji, že jsem se seznámil/a s výsledkem kontroly originality textu závěrečné práce v systému: [-] Theses [-] Turnitin

Komentář k výsledku kontroly:

Kontrola neodhalila žádnou necitovanou shodu s jinými texty.

Celkové hodnocení práce:

Michaela Mendelová předkládá zajímavou práci k aktuálnímu tématu. Je tak škoda, že toto téma v textu dostatečně nezdůvodňuje a zároveň nereflektuje svoje vlastní postoje, ze kterých vychází ve svých interpretacích. I přes tyto výtky doporučuji práci k obhajobě a navrhuji ji hodnotit jako dobrou (E) nebo velmi dobrou (D) v závislosti na průběhu obhajoby.

Datum: 23.1.2023

Podpis: