Výtah kvantitativní obsahové analýzy

Dlouhodobé trendy ve vnímání umělé inteligence veřejností

# Úvod

Lidé si často o umělé inteligenci (v angličtině známé pod pojmem AI) myslí, že to je něco, co nám může škodit, ne-li dokonce ohrozit. Tyto pocity ale nejsou založeny na opravdových vědeckých poznatcích a faktech. Vědci, kteří se zabývají umělou inteligencí se všeobecně shodují, že obavy z umělé inteligence jsou vytvořeny na základě představ, že se inteligentní stroje stanou neovladatelnými a samo rozhodovacími monstry, která zničí lidskou populaci. V dnešní době ale, kdy se umělá inteligence stává lepší a rychlejší než kdy dřív a vysoce inteligentní systémy a technologie jsou levnější, schopnost umělé inteligence pomoct lidstvu ve všech oblastech lidského života, ať už je to v práci, ve škole, v zdravotnictví nebo doma, je čím dál tím větší.

Nemá cenu pokládat si otázku, jestli umělou inteligenci v našich životech chceme. Vysoce inteligentní systémy se staly součástí našich životů už před několika lety a budou s námi čím dál tím více. Měli bychom si ale pokládat otázku, jak se můžeme vzdělávat v oblasti umělé inteligence, jak se přestat bát a jak napomoct tomu, aby inteligentní technologie byly dostupné a bezpečně využité ve prospěch všech lidí. Myslím, že je důležité zajímat se o to, jak lidé umělou inteligenci vidí a sledovat vývoj názorů. Právě proto, jsem si vybrala obsahovou analýzu s názvem *Long-Term Trends in the Public Perception of Artificial Intelligence* (Dlouhodobé trendy ve vnímání umělé inteligence veřejností) od Ethana Fasta a Erica Horvitze.

# Obecné zaměření

Analýzy textů v průběhu času mohou odhalit trendy v zájmu o dané téma. Autoři ve spolupráci se Stanford University použili články z New York Times v rozmezí 30 let k analyzování měnícího se postoje veřejnosti k umělé inteligenci. Pochopení obav veřejnosti z vysoce inteligentních systémů je důležité, protože se tyto obavy mohou promítnout do každodenních činností s potenciálně vážnými dopady. Vědci Fast a Horvitz nejprve definovali sadu indikátorů, které zachycují úroveň zapojení témat o umělé inteligenci a poté sadu indikátorů zobrazující úroveň nadějí nebo obav z umělé inteligence. Tyto ukazatele pak následně využili při zkoumání článků z New York Times. Jejich studie kombinuje techniku crowdsourcingu a anotace. Z každého analyzovaného článku extrahovali všechny zmínky o umělé inteligenci a pomocí placeného crowdsourcingu označili tyto zmínky mírou relevance, pesimismu nebo optimismu vůči umělé inteligenci a přítomností konkrétních nadějí a obav (například „ztráta kontroly nad umělou inteligencí“ nebo „umělá inteligence zlepší zdravotní péči“).

# Výzkumné otázky

Aby zjistili, jak se veřejné vnímání umělé inteligence v průběhu času měnilo, analyzovali již dříve zmiňované sady indikátorů ve článcích publikovaných v New York Times v období od ledna 1986 do června 2016 a zaměřovali se na tyto čtyři výzkumné otázky:

*Otázka 1:*

Jak často je umělá inteligence součástí veřejné diskuse ve srovnání s minulostí?

*Otázka 2:*

Staly se zpravodajské články v průběhu času optimističtější nebo pesimističtější vůči umělé inteligenci?

*Otázka 3:*

Jaké názory se v průběhu času nejvíce spojují s umělou inteligencí?

*Otázka 4:*

Jakých představ se veřejnost v minulosti obávala a již se toho neobává? Jaké nové obavy vyvstaly?

# Metodologie

## Indikátory pocitů z umělé inteligence

Vědci vzali v potaz nejenom intenzitu zapojení témat zabývající se umělou inteligencí, ale i poměr mezi množstvím nadějí a obav z budoucnosti umělé inteligence. Použili následující obecná měřítka; (1) *Angažovanost*: Tento indikátor slouží k měření veřejného zájmu a veřejné angažovanosti v otázce umělé inteligence na základě množství článků věnující se otázkám spojeným s umělou inteligencí v daném období. (2) *Optimismus vs. Pesimismus*: Tento indikátor zachycuje směr diskuse – do jaké míry se diskuse ubírá pozitivnějším nebo pesimističtějším směrem ohledně budoucnosti umělé inteligence. Vědci následně označili naděje jako indikátory pozitivního přístupu (umělá inteligence usnadňuje lidskou práci nebo dokonce úplně osvobozuje od nutnosti pracovat, umělá inteligence zvyšuje zdraví a blahobyt lidí), stejně tak obavy jako indikátory pesimističtějšího přístupu (umělá inteligence ničí různá povolání, umělá inteligence postrádá etické uvažování, což může vést například až ke ztrátám na lidských životech).

## Data: Novinové články z New York Times z období 1986 až 2016

Fast a Horvitz provedli analýzu celé sady článků publikovaných v New York Times v období od ledna 1986 do května 2016 (více než 3 miliony článků). Tento rozsáhlý datový soubor vytvořili ze získaných metadat (název článku, část příspěvku, aktuální URL), které byly přidružené k článkům publikovaným každý jednotlivý den v rámci zkoumaného období. U každého článku pak získali celý text z jeho URL adresy pomocí balíčku BeautifulSoup v Pythonu. Dále chtěli nechat články o umělé inteligenci anotovat. Crowdsourcingové anotace celých zpravodajských článků jsou ale složitý úkol, který vyžaduje velkou časovou náročnost pro pracovníky. Proto články omezili pouze na odstavce. A protože v novinách mají odstavce tendenci být dostatečně samostatné, pracovníci mohli článek přesně zhodnotit, aniž by ho museli přečíst celý.

## Crowdsourcing pro anotaci indikátorů

Ke shromáždění anotací k více než 8 000 odstavcům týkajícím se umělé inteligence vědci použili Amazon Mechanical Turk (AMT). Nejprve seřadili anotace pro postoje k budoucnosti umělé inteligence na pětibodové Likertově škále (od pesimistických po optimistické). Poté shromáždili anotace pro všechny konkrétní naděje a obavy. Vytvořili binární štítky, které označují, zda se v odstavci objevuje naděje nebo znepokojení (např. umělá inteligence bude mít negativní dopad na práci). Nakonec seřadili anotace s vysokou relevancí na pětibodové Likertově škále (od silně nesouvisejících po silně související).

# Výsledky výzkumu

Pomocí shromážděných anotací Fast a Horvitz analyzovali vývoj trendů ve veřejných postojích vůči umělé inteligenci. Analýzu provedli prostřednictvím čtyř výzkumných otázek. Otázku 1 (Jak často je umělá inteligence součástí veřejné diskuse ve srovnání s minulostí?) zodpověděli pomocí indikátoru angažovanosti. Při odpovídání otázky 2 (Staly se zpravodajské články v průběhu času optimističtější nebo pesimističtější vůči umělé inteligenci?) zjistili, že pohled na umělou inteligenci je trvale v průběhu let optimističtější než pesimistický. Také zjistili, že od roku 2009 diskuse o umělé inteligenci spolu s veřejným zájmem o toho téma prudce vzrostla. Kvůli otázce 3 (Jaké názory se v průběhu času nejvíce spojují s umělou inteligencí?) zkoumali klíčová slova nejvíce spojovaná s články o umělé inteligenci. Poslední otázku (Jakých představ se veřejnost v minulosti obávala a již se toho neobává? Jaké nové obavy vyvstaly?) zodpovídali pomocí klíčových slov a škály ukazující stupeň konkrétních nadějí a obav.

Na základě dat Fast a Horvitz zjistili, že umělá inteligence obecně získala v průběhu času větší slovo ve veřejné diskusi. Dále odhalili, že nálada diskuse zůstala v průběhu času optimističtější, i když tento trend není stejný ve všech oblastech (př. dopad umělé inteligence na pracovní místa). Nakonec objevili, že myšlenky jako je obava ze ztráty kontroly nad umělou inteligencí jsou dnes běžnější, než byly v minulosti.

# Závěr

Vědci Fast a Horvitz představili soubor měřítek, která zachycují úroveň angažovanosti, míru pesimismu a optimismu, výskyt konkrétních nadějí a obav a témata, která souvisejí s diskusemi o umělé inteligenci po celá desetiletí. Zjistili, že diskuse o umělé inteligenci od roku 2009 prudce vzrostly a že tyto diskuse jsou více optimističtější než pesimistické. Pokud však zkoumáme konkrétní obavy, zjistíme, že v posledních letech vzrostly obavy ze ztráty kontroly nad vysoce inteligentními systémy, etické obavy a obavy z negativního dopadu umělé inteligence na pracovní místa. Fast a Horvitz také zjistili, že se naděje v potenciál umělé inteligence pomoct ve zdravotnictví a vzdělávání postupem času zvětšují.

Miriam Pogranová

# Zdroje

Fast, E. and Horvitz, E., 2017. Long-term trends in the public perception of artificial intelligence. *AAAI'17: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, [online] pp.963–969. Available at: <https://www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/download/14581/13868> [Accessed 23 October 2020].