

Špela Vintar

Filozofska fakulteta, Ljubljana

DOI: 10.4312/SSJLK.59.21-28

Jezik in umetna inteligenca: kam nas vodijo veliki jezikovni modeli

Prispevek govori o razvoju jezikovne umetne inteligence. V prvem delu govorimo o strojnih prevajalnikih, ki so v široki uporabi že več let, zato nam izkušnje z njimi lahko pomagajo predvideti, kako bo potekalo uvajanje novih jezikovnih modelov v različne dejavnosti. V drugem delu opišemo nekaj primerov uporabe klepetalnega robota ChatGPT v raziskavah in predstavimo najpomembnejše izzive, s katerimi se srečuje družba ob pojavu splošne umetne inteligence.

umetna inteligenca, strojno prevajanje, ChatGPT

This article discusses the development of artificial intelligence. The first part focuses on machine translation, which has already been widely used for several years and can thus help predict how new language models will be introduced into various activities. The second part describes selected examples of applying the ChatGPT chatbot in research and presents the key challenges society must deal with due to the emergence of artificial intelligence in general.

artificial intelligence, machine translation, ChatGPT

1 Uvod

Pred dobrimi petimi leti smo opazili nenaden skok v kakovosti strojnega prevajanja, kot ga je ponujal na primer Google Translate. Če so bila pred letom 2018 strojno prevedena besedila uporabna bolj za grobo razumevanje vsebine in je pot do slovnično in pomensko ustreznega prevoda zahtevala še mnogo človeškega truda, so prevajalniki kar čez noč začeli govoriti skoraj povsem pravilno. Niso imeli več težav s slovenskimi skloni in ujemanjem in tudi daljše povedi so se nena doma brale dokaj tekoče. Kaj se je zgodilo leta 2018?

Danes se z radovednostjo in strahospoštovanjem pogovarjamo s klepetalnimi roboti, ki so naučeni na ogromnih količinah besedil in se lahkotno sprehajajo med različnimi strokovnimi temi, žanri in slogi, medtem ko se odločevalci, znanstveniki in javnost sprašujemo, kje so varne meje (ne)zaupanja v umetno inteligenco in kaj nas čaka v prihodnosti.

V pričujočem prispevku se ukvarjamo s perspektivami, ki jih odpira jezikovna umetna inteligenca, pri tem pa se naslanjamo na dosedanje izkušnje in raziskave s področja strojnega prevajanja. Slednje namreč izhaja iz enake tehnološke podlage, a je v širši uporabi že dovolj časa, da lahko znanstvena in praktična spoznanja do neke mere preslikamo na širše polje intelligentnih jezikovnih tehnologij.

2 Nevronski jezikovni modeli

Kot smo nakazali že v uvodu, so jezikovni modeli z uporabo nevronske mreže povzročili revolucionaren premik na področju jezikovnih tehnologij. Računalniška obdelava jezika je namreč dotlej bolj ali manj sledila ravnem jezikoslovni analizi, kjer se pomikamo od manjših elementov (fonologija, morfologija) prek besed in besednih zvez (leksikologija, frazeologija) do struktur (skladnja) in

celotnih besedil. Za vsako od teh ravni je skrbelo drugo računalniško orodje, poleg naštetih pa je bilo treba razviti še orodja za obdelavo »posebej zahtevnih« jezikovnih prvin: razdvoumljanje večpomenskosti, prepoznavanje imenskih ali terminoloških enot, razreševanje koreferenčnosti in podobno. Večina teh orodij je že pred razmahom nevronskega modelov uporabljala metode strojnega učenja, ki so znanje črpale iz označenih učnih podatkov.

Nevronska mreža je algoritmični pristop za modeliranje poljubnih kompleksnih funkcij. Sestavljena je iz več plasti vozlišč, v katerih poteka procesiranje v obliki optimizacije uteži, pri čemer sta eksplicitni le vhodna in izhodna plast, vmesne pa so skrite. Pomembna značilnost nevronske mreže je povratnost; v fazi učenja se namreč v primeru napačnega izhodnega rezultata algoritem »vrne« nazaj in popravi uteži v vozliščih, ki so reagirala napačno. Globoke nevronske mreže imajo več skritih ravni, danes pa se uporabljajo za izdelavo napovednih modelov na najrazličnejših področjih, od medicinske diagnostike s pomočjo klasifikacije rentgenskih posnetkov, ciljnega trženja na družabnih omrežjih prek analize vedenjskih vzorcev uporabnikov pa vse do finančnih trgov in borz, kjer se na podlagi preteklih gibanj napovedujejo in upravljajo prihodnje transakcije.

Svoje zmožnosti kažejo tudi na področju obdelave jezika. Jezikovni modeli se učijo na velikih količinah besedilnih podatkov, ki jih najdejo na spletu, pri jezikovnih modelih za strojno prevajanje pa so ti podatki pretežno dvojezični. Ker gre za napovedni model, so v istem jezikovnem modelu zakodirani tako slovnični kot pomenski podatki, posamezne besede in njihovi deli pa so v nevronskega modelu predstavljeni z nizom številskih vrednosti, ki jim pravimo vektorska vložitev. V fazi učenja se nevronska mreža uči napovedovati besedne nize tako, da se v učnih podatkih posamezne besede skrijejo, jezikovni model pa svoje napovedi postopno izboljšuje.

Kar je za nadaljnje razumevanje delovanja jezikovnih modelov – kljub tej zelo poenostavljeni razlagi – posebej pomembno, je, da se pri vektorskem modeliranju pomena slednji preslika v nekakšen abstraktni mnogorazsežnostni prostor, kjer se morfološke, semantične in sintagmatske sorodnosti merijo z razdaljami med vektorji. Jezikovni model zato pri svojih predikcijah, pa naj gre za prevajanje, generiranje besedila ali kaj tretjega, ne uporablja naučenih verjetnosti posameznih besednih nizov, ampak abstraktne vektorske vložitve, ki simulirajo nekakšen konceptualni prostor.

V času pisanja tega prispevka (marec 2023) je v raziskovalni skupnosti znanih več deset velikih jezikovnih modelov, od katerih so nekateri dostopni tudi širši javnosti (ChatGPT, GPT-4). Razlikujejo se predvsem po količini učnih podatkov in številu parametrov; GPT-4 naj bi bil zgrajen s kar bilijonom parametrov. Sočasno z bliskovitim napredkom velikih jezikovnih modelov se v javnosti in strokovnih skupnostih množijo najrazličnejši odzivi, ki jih na tem mestu ne bomo povzemali. Morda pa je vredno omeniti le poziv inštituta Future of Life po vsaj šestmesečnem moratoriju na nadaljevanje razvoja vse večjih in večjih jezikovnih modelov, ki ga je podpisalo impresivno visoko število globalnih tehnominenc, od Elona Muska, Steva Wozniaka in Yuvala Noaha Hararija dalje.¹ Podpisani so namreč prepričani, da se z razvojem splošne umetne inteligence človeštvo podaja na pot brez povratka, na soočenje s tako prelomnim civilizacijskim preobratom pa niti približno nismo pripravljeni.

V nadaljevanju se posvetimo tistim vidikom jezikovnih modelov, ki so doslej pokazali preverljivo uporabno vrednost, in razpravljamo o njihovih možnih učinkih v prihodnosti.

1 <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

3 Strojno prevajanje

Kot smo nakazali že v uvodu, postajajo strojni prevajalniki vse boljši, še posebej od prehoda na nevronske jezikovne modele. Ti so danes uporabljeni pri vseh velikih spletnih prevajalnikih, kot so DeepL, Google Translate, Microsoft Bing in Yandex. Na svetovni konferenci o strojnem prevajanju 2018 (WMT2018) se je prvič zgodilo, da sta dva prevajalnika pri prevajanju novic za jezikovna para kitajščina-angleščina in angleščina-češčina (Bojar idr. 2018) dosegla človeško raven, slednji jo je na omenjenem tekmovanju celo presegel. Leto kasneje je bilo sistemov, ki so dosegali podobno kakovost kot človeški prevajalci, a le za omejeno področje prevajanja novic, že prek deset (Barrault idr. 2019). Lani so na tej konferenci področje preverjanja uspešnosti razširili na različne žanre splošnih besedil, poleg človeškega ocenjevanja kakovosti posameznih prevodov pa so izvedli tudi podrobno analizo napak. Čeprav je pri strojnih prevodih število napak v povprečju presegallo število napak v referenčnih prevodih (ne pozabimo, da smo tudi ljudje zmotljivi), je nekaj prevajalnikov tudi po tem kriteriju presegllo kakovost študentskih prevodov (Kocmi idr. 2022).

V zadnjem času je bilo izvedenih še nekaj zanimivih raziskav, ki kažejo, da strojno prevajanje postaja vse boljše tudi za manjše jezikovne pare in za nestandardna besedila. Campo Bayón in Sánchez-Gijón (2022) sta razvili strojni prevajalnik tvtov iz španščine v galicijsčino, nato pa pri uporabnikih preverjali, kako dobro razlikujejo med strojno prevedenimi in izvirnimi tviti. Presenetljiva ugotovitev je bila, da so uporabniki lažje prepoznavali strojne prevode kratkih tvtov, pri daljših tvitih pa so strojni prevodi uporabnike zlahka preslepili, da so bili izvorno spisani v galicijsčini. Očitno je na tvitih naučeni jezikovni model dobro posnemal pogovorne značilnosti galicijskih tviterašev, ampak le, kadar je bilo besedila nekoliko več.

Vse večje sprejemanje strojnega prevajanja (angl. *MT acceptance*) kot koristne tehnologije je neizbežno povzročilo tektonske premike v jezikovni industriji, katere pomemben delež predstavlja jo storitve prevajanja in tolmačenja. Orodja za računalniško podprto prevajanje, kot so RWS Trados, memoQ, Phrase ipd., brez izjeme ponujajo integracijo strojnih prevajalnikov v delovno okolje, in sicer tako da imamo med prevajanjem v prevajalskem programu na voljo prevodne zadetke iz lastnih pomnilniških baz, javnih pomnilnikov prevodov in poljubnih strojnih prevajalnikov. Quintana in Castilho (2022) v svoji primerjavi prevajalskih orodij glede integracije strojnega prevajanja ugotavljata precejšnje razlike, videti pa je, da pri prevajalcih narašča priljubljenost orodij v oblaku in da jim postaja udobje glede stopnje in načina integracije pomemben dejavnik pri izbiri orodja.

Ali strojno prevajanje pomeni konec človeških prevajalcev? Vprašanje se zdi posebej relevantno v času, ko se ves svet ukvarja s sprejemanjem ali nesprejemanjem klepetalnih robotov ter groženj, ki jih ti predstavljajo nekaterim človeškim dejavnostim. Na področju prevajanja lahko namreč razvoj in odzive nanj spremljamo že dlje časa, saj se je uvajanje strojnih prevajalnikov v profesionalna okolja začelo že pred vsaj desetimi leti. Že takrat so mnogi prevajalci, prevodni teoretiki, pa tudi učitelji prevajanja zavzeli stališče, da so strojni prevajalniki nekakšno nujno zlo, ki obstaja onkraj univerzuma »pravega« prevajanja, pogosto pa je bilo slišati mnenja (ta se pojavljajo še danes), da »so strojni prevajalniki sicer vse boljši, a *nikoli* ne bodo dosegli človeške kakovosti«. Mnenja se običajno podkrepijo z razlogi, kaj vse računalnikom manjka, da bi se lahko kdaj merili z ljudmi.

Zgodnja leta uvajanja prevajalnikov v profesionalna prevajalska okolja so bila zaznamovana s številnimi študijami produktivnosti, poudarjanjem kognitivnih razlik med prevajanjem in popravljanjem strojnih prevodov (angl. *post-editing*) in z analizami napak, ki jih je bilo takrat še v izobilju. Odkar so na voljo nevronske prevajalniki, je slovničnih napak bistveno manj, popravljanje pa je postalo po mnenju nekaterih še težje, saj se napake skrivajo v brezhibno oblikovanih povedih in v slogovno nadvse prepričljivih ubeseditvah. Novejše raziskave kažejo, da se produktivnost prevajalcev z uporabo strojnih prevajalnikov poveča za kar štirikrat (Murgolo 2022). V omenjeni študiji je avtorica primerjala število besed, ki jih prevajalci ali popravljalci ustvarijo v določeni časovni enoti, in sicer v treh scenarijih: pri prevajanju brez pomnilnikov prevodov ali strojnih prevajalnikov, pri popravljanju delno prevedenega besedila z zadetki iz pomnilnika prevodov in pri popravljanju strojno prevedenega besedila. Kljub temu da so imeli prevajalci pri drugem scenariju na voljo pomnilnik prevodov z vsaj 85-odstotno podobnimi zadetki, so dosegli daleč največjo produktivnost pri popravljanju strojno prevedenega besedila.

Zanimive so tudi izkušnje največje prevajalske agencije v Evropi, Generalnega direktorata za prevajanje Evropske komisije (DGT), kjer je strojno prevajanje v nekaj letih marsikje povsem nadomestilo človeško delo. Evropska komisija je z zajetnim finančnim vložkom zgradila prevajalnik eTranslation, ki se po kakovosti danes zlahka kosa z največjimi, na nekaterih uradnih spletiščih Unije pa se prevodi v jezike članic zagotavljajo samo še s strojnim prevajalnikom. Razumljivo je, da prevajalci novih tehnologij ne sprejemajo brez strahov in skrbi. V raziskavi sprejemanja strojnega prevajanja, ki sta jo leta 2017 med prevajalci DGT izvedla Rossi in Chevrot (2019), je več kot 40 odstotkov vprašanih oseb strojno prevajanje izpostavilo kot grožnjo prevajalskemu poklicu, čeprav je bila raziskava izvedena še v obdobju pred uvedbo nevronskega sistema eTranslation. Njegov statistični predhodnik MT@EC je bil namreč po kakovosti precej slabši in bi zanj pričakovali, da bo pri prevajalcih vzbujal manj strahov. Raziskava je pokazala tudi pomembno korelacijo med slabim poznavanjem tehnologije strojnega prevajanja in občutkom ogroženosti, z drugimi besedami so tisti prevajalci, ki so svoje poznavanje tehnologije ocenili kot šibko ali nezadostno, izrazili tudi največ skrbi in strahov v zvezi s prihodnostjo.

Da bo podoba popolnejša, omenimo še vsakoletno raziskavo stanja v evropski jezikovni industriji (ELIS 2022), ki se izvaja vse od leta 2013 in zajema reprezentativen vzorec akterjev na tem področju. Zadnje poročilo, objavljeno marca 2023, izpostavlja veliko spremembo v odnosu do novih tehnologij, saj kar 70 odstotkov sodelujočih redno ali občasno uporablja strojne prevajalnike. Če so bili leta 2013 občutki udeležencev glede tehnološkega razvoja še pretežno negativni, pa je lani kar 65 odstotkov sodelujočih podjetij menilo, da so izboljšave nevronske prevajalnike priložnost in ne grožnja. Res pa je, da takšnega mnenja niso delili udeleženci, ki se preživljajo kot samostojni podjetniki, niti predstavniki izobraževalnih ustanov – pri obojih je prevladoval občutek, da je prevajalski poklic v prihodnosti ogrožen.

Kot vidimo iz prejšnjih odstavkov, so se strojni prevajalniki v zgolj desetletju iz obskurnih, nerodnih in negotovost vzbujajočih orodij prelevili v zanesljiv in široko uporaben pripomoček, ki bistveno zvišuje produktivnost, znižuje stroške in mnogim jezikovno neveščim uporabnikom omogoča dostop do tujejezičnih vsebin, pa tudi tujejezično komunikacijo. Po drugi strani se je prevajalski poklic zaradi tehnologij korenito spremenil, in bo po vsej verjetnosti zanj v prihodnje vse manj

zanimanja. Na njegovo mesto utegnejo priti drugačni profili jezikovnih profesionalcev, ki bodo morali imeti gibčne jezikovne veščine, ob tem pa še široko jezikovnotehnološko podlago (Vintar 2019).

4 Jezikovni modeli v praksi

V nasprotju s prevajalniki, ki imajo za sabo okrog 70 let razvojne zgodovine, so jezikovni modeli razmeroma mlada tehnologija, z velikimi jezikovnimi modeli pa se z uporabniškega vidika srečujemo šele nekaj mesecev. Kljub temu sta ChatGPT, kmalu zatem pa njegov naslednik GPT-4, sprožila pravo kulturno revolucijo, ki odmeva v znanstvenih krogih, izobraževanju, medijih in v širši javnosti. Le štiri mesece po splostitvi klepetalnega robota ChatGPT na portalu Google Scholar najdemo že prek 7000 zapisov o njem, od tega sicer večino komentatorskih besedil iz strokovnega tiska, pa tudi uvodnikov in pisem bralcev v ugledne znanstvene revije. Nekatere prispevke je spisal kar sam robot, kot priznavajo avtorji v uvodu, z namenom demonstracije njegovih zmožnosti in šibkosti pri uporabi v znanstvene namene.

Naštejmo nekaj področij, kjer se od tovrstnih tehnologij pričakuje največji učinek. Prvo je brez dvoma medicina, saj se umetna inteligenca lahko uporablja za hitrejšo in natančnejšo diagnostiko, posledično pa tudi uspešnejše zdravljenje in znižanje stroškov. Antaki idr. (2023) so preverjali uporabo ChatGPT-ja na področju oftalmologije in pokazali, da robot pravilno prepoznava oftalmološka obolenja in ustvarja načrte zdravljenja, ki so v skladu s strokovnimi smernicami. Jeblick idr. (2023) so ChatGPT vgradili v jezikovno aplikacijo, ki poenostavlja radiološke izvide. Raziskava je pokazala, da sistem dosega visoko natančnost tako pri vsebinski točnosti poenostavljenih izvidov kot pri slovnični in slogovni ustreznosti. Aplikacija je posebej pomembna za bolnike, ki so s pomočjo berljivejših izvidov lažje razumeli svojo zdravstveno situacijo in se odločali o nadaljnjih posegih. Tudi številni drugi avtorji razpravljajo o potencialih, ki jih ChatGPT prinaša v različne medicinske vede, na primer v kirurgijo, stomatologijo in zdravniško oskrbo, koristen pa je tudi pri pregledovanju medicinske literature in kot pripomoček pri medicinskem pisanju.

Naslednje področje je, pričakovano, računalniška obdelava jezika. ChatGPT se, ob ustreznih iztočnicah (angl. *prompts*), dobro odreže pri razdvoumljanju večpomenskih besed (Ortega-Martin idr. 2023), povzemanju besedil (Yang idr. 2023) in strojnem prevajanju (Jiao idr. 2023), o njegovih jezikovnih zmožnostih pa poleg objavljenih znanstvenih člankov govori morje primerov, zapisov in objav na družabnih omrežjih, v splošnih medijih in na temu posebej namenjenem kanalu na Discordu.

Prav zaradi jezikovnih zmožnosti so se na pojav ChatGPT-ja med prvimi odzvale univerze, saj je univerzitetni esej na številnih vodilnih svetovnih univerzah ključno orodje za ocenjevanje študentov, skoraj pri vseh študijih pa je predpisano tudi zaključno delo. Ker pisnih izdelkov ChatGPT-ja sistemi za preverjanje pristnosti ne zaznajo kot plagiat (saj to v resnici ni), je med univerzami vzvalovila skrb, da bo uveljavljene načine preverjanja znanja treba korenito spremeniti (Susnjak 2022). Vendar se hkrati rojevajo tudi številne zamisli, kako inteligentne robote vključiti v izobraževanje in si z njimi pomagati. Frieder idr. (2023) na primer opisujejo preskus uporabe ChatGPT-ja pri reševanju matematičnih nalog in ga primerjajo z drugimi modeli, Chen (2023) pa poglobljeno razpravlja o koristih in možnostih transformacije bibliotekarskih storitev.

Čeprav bo pričujoči prispevek ob objavi že precej zastarel, saj se novi znanstveni in strokovni prispevki o tej temi objavljajo vsakodnevno, je v času pisanja eden boljših pregledov znanstvene literature na temo ChatGPT-ja članek v predobjavi Koubaa idr. (2023).

In pri nas? Najprej smo lahko ponosni, da imamo v Sloveniji vrhunske in svetovno znane strokovnjake za umetno inteligenco, računalniško obdelavo jezika in povezane tehnologije. Pojav ChatGPT-ja je tudi v našem okolju spodbudil novinarje, komunikatorje znanosti, digitalne navdušence in celo filozofe k pisanju o njem. Nekoliko resnejših preskusov te tehnologije za slovenščino je manj, omenimo pa lahko študijo s področja socialnega dela, ki primerja človeško in umetno inteligenco pri izvedbi kvalitativne analize (Mesec 2023). Omenjeni prispevek nakazuje, da se jezikovni model dobro odreže tudi pri iztočnicah v slovenščini, saj avtor ugotavlja, da robot ponuja presenetljivo ustrezne odgovore. V sklepnem delu avtor zapiše še:

Kakšno koli omalovaževanje umetne inteligence je neprimerno. Program ni popoln, je pa blizu temu, sodeč po rezultatih. [...]

Sposoben je ustvarjalnega parafraziranja, to je, povzemanja izjav, tako da izrazi njihovo bistvo, lahko bi rekli celo, da je v tem pomenu sposoben »brati med vrsticami«, to je, izreči, kar ni bilo izrečeno, bilo pa je mišljeno. Tega marsikatero človeško bitje ne zna. (Mesec 2023)

5 Ugotovitve

Skoraj vsa doslej citirana dela navajajo tudi številne zadržke v zvezi z uporabo inteligentnih tehnologij v praksi. Med najpogostejše omenjenimi izzivi so:

- **varstvo podatkov:** ChatGPT in sorodni jezikovni modeli se učijo iz ogromnih količin besedil, ki pogosto vsebujejo osebne podatke o uporabnikih (npr. dnevniki spletnih klepetov). Zaenkrat je zakonska regulativa na tem področju bistveno preohlapna.
- **etika uporabe:** Pri tehnologiji, kakršna je ChatGPT, so možnosti zlorab precejšnje, čeprav se razvijalci trudijo zlonamerneže kar najbolj omejiti. Ne gre le za zlorabe v smislu (lažnega) avtorstva člankov, esejev in diplomskih del, ampak tudi za grožnjo namernega zavajanja in manipulacije prek lažnih novic, spletnih komunikacij, objav na družbenih omrežjih in podobno. Če pa v etično razmišljanje vključimo še inteligentne tehnologije jutrišnjega dne, se znajdemo na še bistveno bolj spolzkem in na trenutke distopičnem terenu, kjer se lahko upravičeno sprašujemo o tem, ali (oziroma bolje kdaj) se bo umetna inteligenca obrnila proti človeštvu in mu začela škoditi. S takšnimi in sorodnimi vprašanji se ukvarja več *miselnih tankov* po svetu, na primer že omenjeni inštitut Future of Life in Institute for the Future of Humanity pri oxfordski univerzi, za usklajevanje človeških interesov z morebitnimi interesi umetne inteligence pa se je uveljavil izraz poravnava UI (angl. *AI alignment*).
- **pristranskost:** Obstaja velika (in večkrat izpričana) nevarnost, da klepetalni roboti in druge inteligentne tehnologije delujejo diskriminatorno do določenih deprivilegiranih skupin ter da utrjujejo stereotipe in stigmatizacijo. Ta izziv je mogoče rešiti s strožjim nadzorom nad učnimi podatki, kar pa je ob sedanjih in prihodnjih količinah skoraj nemogoče izvajati v praksi.

- **enakopravnost:** S tem mislimo predvsem na enakopravnost pri dostopanju do inteligentnih storitev in njihovi uporabi, saj nam v nasprotnem primeru grozi še večja razslojitev sveta in družb na tehnološko podprte elite in nemočne preostale.
- **transparentnost:** ChatGPT in sorodniki so nevronske napovedni modeli, ki svojih napovedi, se pravi generiranih besedil, ne pojasnjujejo, prav tako tudi njihovi razvijalci težko pojasnijo razloge, da je v določenem primeru prišlo do določene predikcije. S tega vidika je njihovo delovanje netransparentno, to pa je še posebej v primeru zlorab in pri vprašanih etike in odgovornosti problematičen vidik.
- **zanesljivost:** S tem mislimo na dejstvo, da se ChatGPT in njegovi nasledniki še vedno motijo, izmišljujejo in navajajo neresničnosti, ki jih celo prepričljivo utemeljujejo, za kar se je prijel izraz »halucinacije«. Seveda se razvijalci s tem vidikom intenzivno ukvarjajo, a še bolj pomembno je izobraževanje splošne javnosti o delovanju jezikovnih modelov in z njim povezanimi tveganji.

Kljub peticiji Elona Muska in sopodpisnikov lahko pričakujemo, da bodo vse bolj inteligentni roboti vse bolj sooblikovali naše vsakdanje življenje, v dobrem in slabem smislu. Neizbežno se bodo vse bolj uporabljali v različnih profesionalnih okoljih in pri tem – kot kažejo izkušnje s sveta prevajalstva – naleteli na mnogo negativnih odzivov. Po drugi strani so najbolj podcenjujoča in zavračajoča stališča pogosto plod nepoznavanja in strahu pred lastnimi šibkostmi. Ker živimo v vznemirljivem času, ko je prostora za konstruktivne ideje še dovolj, prav tako pa potrebujemo čim več tehničnih razmislekov o priložnostih in tveganjih, ki jih jezikovni modeli prinašajo, je ključna radovednost. Slednje smo skušali spodbuditi tudi s tem prispevkom.

Literatura

- ANTAKI, Fares, TOUMA, Samir, MILAD, Daniel, EL-KHOURY, Jonathan, DUVAL, Renaud, 2023: Evaluating the performance of ChatGPT in ophthalmology: An analysis of its successes and shortcomings. *medRxiv* 2023/1.
- BARRAULT, Loic, BOJAR, Ondřej, COSTA-JUSSA, Marta R., FEDERMANN, Christian, FISHEL, Mark, GRAHAM, Yvette, HADDOW, Barry idr., 2019: Findings of the 2019 conference on machine translation (WMT19). *Association for Computational Linguistics*.
- BOJAR, Ondřej, FEDERMANN, Christian, FISHEL, Mark, GRAHAM, Yvette, HADDOW, Barry, HUCK, Matthias, KOEHN, Philipp, MONZ, Christof, 2018: Findings of the 2018 Conference on Machine Translation (WMT18). *Proceedings of the Third Conference on Machine Translation. Association for Computational Linguistics*. 272–307.
- do CAMPO BAYÓN, María, SÁNCHEZ-GUJÓN, Pilar, 2022: Evaluating NMT: superior, inferior, or equivalent to texts originally written by humans. *New Trends in Translation and Technology NeTTT 2022, 4–6 July 2022*. Rhodes Island, Greece. 27–36.
- CARO QUINTANA, Rocio, CASTILHO, Sheila, 2022: A review of the Integration of Machine Translation in CAT tools. *New Trends in Translation and Technology 2022 NeTTT 2022, 4–6 July 2022*. Rhodes Island, Greece. 214–221.
- CHEN, Xiaotian, 2023: ChatGPT and its possible impact on library reference services. *Internet Reference Services Quarterly*. 1–9.
- European Language Industry Survey, 2022: Trends, expectations and concerns of the European language industry. ELIA: https://fit-europe-rc.org/wp-content/uploads/2022/03/ELIS-2022_survey_results_final_report.pdf?x85225
- FRIEDER, Simon, PINCHETTI, Luca, GRIFFITHS, Ryan-Rhys, SALVATORI, Tommaso, LUKASIEWICZ, Thomas, PETERSEN, Philip Christian, CHEVALIER, Alexis, BERNER, Julius, 2023: Mathematical capabilities of ChatGPT. arXiv preprint: 2301.13867.
- JEBLICK, Katharina, SCHACHTNER, Balthasar, DEXL, Jakob, MITTERMEIER, Andreas, STÜBER, Anna Theresa, TOPALIS, Johann, WEBER, Thomas, WESP, Philipp, SABEL, Bastian, RICKE, Jens, idr., 2022: Chatgpt makes medicine easy to swallow: An exploratory case study on simplified radiology reports. arXiv:2212.14882. <https://doi.org/10.48550/arXiv.2212.14882>

- JIAO, Wenxiang, WANG, Wenxuan, HUANG, Jen-tse, WANG, Xing, TU, Zhaopeng, 2023: Is ChatGPT a good translator? A preliminary study. arXiv: 2301.08745. <https://doi.org/10.48550/arXiv.2301.08745>
- KOCMI, Tom, BAWDEN, Rachel, BOJAR, Ondřej, DVORKOVICH, Anton, FEDERMANN, Christian, FISHEL, Mark, GOWDA, Thamme, idr., 2022: Findings of the 2022 conference on machine translation (WMT22). *Proceedings of the Seventh Conference on Machine Translation (WMT)*. 1–45.
- KOUBAA, Anis, BOULILA, Wadii, GHOUTI, Lahouari, ALZAHM, Ayyub, LATIF, Shahid, 2023: Exploring ChatGPT Capabilities and Limitations: A Critical Review of the NLP Game Changer. Preprints.org, 2023030438. <https://doi.org/10.20944/preprints202303.0438.v1>
- MESEC, Blaž, 2023: Jezikovni model umetne inteligence chatGPT – pripomoček kvalitativne analize. V recenziji pri reviji *Socialno delo*, vir članka osebna komunikacija z avtorjem.
- MURGOLO, Elena, 2022: Productivity Evaluation in MT Post-Editing and Fuzzy Matches Ed-iting. Setting the Threshold. *New Trends in Translation and Technology NeTTT 2022. 4-6 July 2022*. Rhodes Island, Greece. 1–8.
- ORTEGA-MARTÍN, Miguel, GARCÍA-SIERRA, Oscar, ARDOIZ, Alfonso, ÁLVAREZ, Jorge, ARMENTEROS, Juan Carlos, ALONSO, Armenteros, 2023: Linguistic ambiguity analysis in ChatGPT. arXiv:2302.06426. <https://doi.org/10.48550/arXiv.2302.06426>
- ROSSI, Caroline, CHEVROT, Jean-Pierre, 2019: Uses and perceptions of Machine Translation at the European Commission. *Journal of specialised translation (JoSTrans)*, halshs-01893120v2.
- SUSNJAK, Teo, 2022: ChatGPT: The end of online exam integrity? arXiv: 2212.09292. <https://doi.org/10.48550/arXiv.2212.09292>
- VINTAR, Špela, 2019: Language in the Age of Dataism. *Slovenščina 2.0: empirical, applied and interdisciplinary research* 7/1. 126–143.
- YANG, Xianjun, LI, Yan, ZHANG, Xinlu, CHEN, Haifeng, CHENG, Wei, 2023: Exploring the limits of ChatGPT for query or aspect-based text summarization. arXiv: 2302.08081. <https://doi.org/10.48550/arXiv.2302.08081>