

Návrh Stanoviska

Vedeckej rady a pedagógov VŠ ZaSP sv. Alžbety v Bratislave, n.o., ku zapojenie AI do vzdelávania a hodnotenia

O nástroje umelej inteligencie (AI) ako ChatGPT, DALLÉ-2, CoPilot a najnovšie aj Google Bard je veľký záujem a experimentuje sa s nimi v rôznych oblastiach, vrátane akademického prostredia a vzdelávania.

Všetky tieto prístupy môžu byť užitočnými nástrojmi na generovanie obsahu, ktorý môže prispieť k hodnotenej práci, čo viedlo niektoré inštitúcie k zákazu ich používania.

Veríme, že tieto nástroje sú potenciálne transformačné a zároveň nesú so sebou aj veľkú mieru rizika, tým, že sa budú používať (alebo sa už používajú) na mnohých odborných pracoviskách, a že **namiesto toho, aby sme sa snažili zakázať vám ich používanie, vás, ako študentov VŠ ZaSP sv. Alžbety v Bratislave podporujeme v ich efektívnom, etickom a transparentnom používaní.**

Aj keď sú atraktívne výkonné a ľahko použiteľné, **môžu poskytovať aj zavádzajúce alebo nesprávne informácie a môžu negatívne ovplyvniť vaše vzdelávanie a obsah osvojených vedomostí**, pretože môžu ponúkať skratky, ktoré **znižujú potrebu kritického myslenia**, ktoré je kľúčom k hlbokému a zmysluplnému učeniu. Musíte si byť vedomí aj rozdielu medzi primeraným používaním takýchto nástrojov a tým, v akom bode by sa ich použitie mohlo považovať za poskytnutie tzv. „**nespravodlivej výhody**“ (poznámka: tento termín sa používa v súčasnosti pre označenie foriem plagiátorstva, ktoré majú špecifické zameranie na proces získavania obsahu, t.j. produkty umelej inteligencie).

Je dôležité, aby ste ako študenti VŠ ZaSP sv. Alžbety v Bratislave na generovanie svojich písomných, grafických alebo iných výstupov **nepoužívali** nástroje AI a neodoslali vaše práce, ako keby to bola vaša vlastná práca. V tomto stanovisku pedagógov VŠ ZaSP sv. Alžbety v Bratislave vysvetľujeme, **ako by sa AI mohla používať a ako priznať jej použitie.**

Čo je umelá inteligencia?

Úvod do umelej inteligencie (AI)

Umelá inteligencia (AI) je zastrešujúci pojem pre technológiu schopnú inteligentne vykonávať úlohy. Priekopníci, odborníci, výskumníci, spoločnosti, organizácie, médiá a tak ďalej navrhli rôzne definície. Niektoré z týchto definícií zahŕňajú:

"AI je schopnosť stroja zobrazovať ľudské schopnosti, ako je uvažovanie, učenie, plánovanie a kreativita." ([Európsky parlament, 2021](#))

"Schopnosť digitálneho počítača alebo počítačom riadeného robota vykonávať úlohy bežne spojené s intelligentnými bytosťami." ([Encyklopédia Britannica, 2023](#))

Jednoducho povedané, AI možno chápať ako mozog pre počítače, roboty a ďalšie inteligentné technológie. Rovnako ako sa študenti učia získať lepšie známky na skúškach, AI sa môže naučiť zlepšovať sa v rôznych úlohách. Spočiatku to nemusí byť veľmi dobré v konkrétnej úlohe, opäť ako keď sa dieťa najprv učí niečo nové. Ale keďže sa AI neustále učí z údajov a skúseností, jej výkon sa časom zlepšuje.

¹ Poznámka: podľa poslednej konzultácie s OpenAI, ktorej obsah bol aktualizovaný naposledy ku septembru 2021, tento typ AI je stále vo vývoji a ešte neexistuje.

AI možno rozdeliť do dvoch hlavných typov:

Slabá AI: Tento typ AI je navrhnutý na vykonávanie úzko vymedzenej úlohy (napr. rozpoznávanie tváre alebo vyhľadávanie na internete) a funguje v obmedzenom kontexte. Napriek svojim schopnostiam nemá skutočnú inteligenciu, ani sebauvedomenie. Príkladmi tohto druhu AI sú napr. Siri, Google Assistant, Amazon Alexa a mnoho ďalších systémov poháňaných AI.

Všeobecná alebo silná AI: Tento typ AI, odkazuje na systémy riešiace úlohy rovnako dobre alebo dokonca lepšie ako človek a to bez nutnosti predchádzajúceho určenia jednotlivých úzko vymedzených úloh. Dokáže porozumieť, učiť sa, prispôbovať a implementovať vedomosti v širokom spektre úloh, pričom preukazuje inteligenciu rovnakú alebo prevyšujúcu ľudskú inteligenciu. Je to často typ AI zobrazený v sci-fi, (ako napríklad HAL 9000 z roku 2001: Vesmírna odysea alebo Dáta zo Star Treku)¹.

„Technológie AI sa čoraz viac integrujú do rôznych odvetví a aspektov života, od zdravotníctva a vzdelávania až po dopravu a zábavu. Hoci má AI množstvo výhod, vyvoláva aj etické a spoločenské obavy, ako je automatizácia práce, problémy s ochranou súkromia a možnosť zneužitia.“ (OpenAI, 2024, 15. február)

Aplikácie a obmedzenia

Vzhľadom na dynamickú povahu strojového učenia sa nástroje založené na AI neustále vyvíjajú a zdokonaľujú, čo vedie k rýchlym zmenám v tejto oblasti. Preto je nasledujúca diskusia kontextovo obmedzená na čas, v ktorom bola táto správa napísaná (ku 20.2.2024).

Na čo je AI dobrá?

Nástroje využívajúce AI nám môžu pomôcť mnohými spôsobmi, napríklad:

Odpovedanie na otázky, na ktoré sú odpovede dostupné na internete alebo ku ktorým sa možno dostať prostredníctvom verejného odkazu, ktorý používateľ poskytne; napr.: "Akú farbu má jablko?"

Preskúmanie a zhrnutie písomných materiálov – to môže byť užitočné najmä pri zostavovaní prehľadu opisnej literatúry alebo abstraktu.

Pomáha pri hľadaní a filtrovaní informácií pre rôzne aplikácie, od každodenných životných situácií, ako je napríklad učenie sa hry, až po akademické účely, ako je vyhľadávanie relevantnej výskumnej literatúry.

Plánovanie a štruktúra písomných materiálov.

Pôsobenie ako tzv. „digitálni osobní asistenti“: napríklad Siri (vstavaná, hlasom ovládaná virtuálna asistentka dostupná pre používateľov Apple) môže byť použitá na pomoc pri organizovaní každodenných rutín.

Poskytovanie strojových prekladov.

Navrhovanie, generovanie nápadov a brainstorming.

Vytváranie grafov, obrázkov a vizuálov na základe poskytnutých informácií.

Pomáha zlepšovať štýl a štruktúru písania atď.

¹ Poznámka: podľa poslednej konzultácie s OpenAI, ktorej obsah bol aktualizovaný naposledy ku septembru 2021, tento typ AI je stále vo vývoji a ešte neexistuje.

Preskúmanie a kritická analýza písomných materiálov s cieľom posúdiť ich platnosť.

Pomáha zlepšiť gramatiku a štruktúru písania – obzvlášť užitočná môže byť, ak je angličtina vaším druhým jazykom.

Experimentovanie s rôznymi štýlmi písania.

Obmedzenia

Umelá a ľudská inteligencia však nie sú to isté. Nástroje AI nerozumejú ničomu, čo produkujú, ani nerozumejú tomu, čo slová, ktoré vytvárajú znamenajú, keď sa aplikujú na skutočný svet (a do rôznych situácií a kontextov).

OpenAI, tvorcovia ChatGPT, poskytli užitočné rady pre pedagógov a študentov – v skratke:

Aj keď sa ich výstup môže javiť ako hodnoverný a dobre napísaný, nástroje AI sa často pokazia (napr. preťažia pod množstvom požiadaviek rôznych užívateľov prihlásených v jednom čase) a nemožno sa na ne spoľahnúť, pokiaľ ide o faktickú presnosť.

Dosahujú lepšie výsledky v predmetoch, o ktorých sa často píše, a horšie v špecializovaných oblastiach a najmä tam, kde je potrebná aktualizácia dôkazov z vedy za posledné 2 roky (viď odkaz 1).

V súčasnosti nemôžu poskytnúť referencie – vyrábajú dobre formátované, ale fiktívne citácie.

Môžu udržiavať stereotypy, predsudky a nereálne perspektívy.

Čo je ešte zásadnejšie, prílišné spoliehanie sa na tieto nástroje zníži vaše príležitosti na zdokonaľovanie schopností písania, kritického myslenia a hodnotenia, teda kľúčových akademických a profesionálnych zručností, ktoré budete potrebovať, keď budete napredovať v štúdiu a ďalej.

Čo však tieto nástroje môžu poskytnúť, je príležitosť vybudovať si kritické analytické a hodnotiace schopnosti – možno štúdiom a kritizovaním toho, čo vyprodukujú, a úsudkom o tom, či to, čo vyprodukujú, je skutočne platné a vierohodné.

Použitie AI a nespravodlivá výhoda na akademickej pôde

Používanie nástrojov AI na pomoc s takými aktivitami, ako je generovanie nápadov alebo vaše plánovanie napr. výskumu, môže byť vhodným využitím, hoci je potrebné zvážiť váš kontext a povahu hodnotenia a skúmanej problematiky. Nie je prijateľné používať tieto nástroje na napísanie celej práce (vrátane záverečných Bc. Mgr., PhD., PhD. Prác) od začiatku do konca. Majte tiež na pamäti, že slová a nápady generované niektorými nástrojmi AI využívajú nápady iných ľudských autorov bez toho, aby na ne odkazovali, čo je v súčasnej dobe samo osebe kontroverzné a je široká zhoda na tom, že takýto čin je považovaný za formu plagiátorstva.

Vaši učitelia vám budú môcť poskytnúť konkrétnejšie pokyny o tom, aké nástroje, ak nejaké, sú povolené v kontexte vašich vlastných úloh z príslušného predmetu alebo pri písaní istého druhu prác, ale princípy uvedené ďalej platia pre všetky hodnotenia prác na VŠ ZaSP sv. Alžbety v Bratislave a jej všetkých študijných programoch.

¹ Poznámka: podľa poslednej konzultácie s OpenAI, ktorej obsah bol aktualizovaný naposledy ku septembru 2021, tento typ AI je stále vo vývoji a ešte neexistuje.

Nespravodlivou výhodou je:

Keď niečo skrývate a nie ste transparentní vo svojom prístupe (napr. ku písaniu, generovaniu nápadov a pod.).

Keď niečo beriete priamo od niekoho (alebo niečoho) iného a neodvolávate sa, ani nepriznáвате takého použitie.

Keď niečo nepoužívate vhodným spôsobom.

Keď sa neriadite pokynmi na hodnotenie od svojho pedagóga.

Ak majú pedagógovia podozrenie, že sa snažíte vydávať výstupy generované AI za svoju vlastnú prácu, platia predpisy v súlade so smernicou o plagiátorstve a v takomto prípade môžete byť pozvaní na prešetrenie písania samotnej práce, dokladovanie zdrojovej literatúry a pod., aby sa zistilo autorstvo vášho diela.

Priznanie používania AI a citovanie AI

Toto usmernenie/ stanovisko predstavuje všeobecný prístup VŠ ZaSP sv. Alžbety v Bratislave k priznaniu používania generatívnej AI a citovaniu generatívnej AI. Generatívna AI sa rýchlo vyvíja a zatiaľ neexistuje všeobecný konsenzus o tom, ako ju priznať a odkázať sa na ňu pri všetkých možných variantoch autorského diela. Toto usmernenie/ stanovisko sa preto bude naďalej revidovať a aktualizovať a vznikne k nemu aj smernica, ktorá bude ošetrovať túto problematiku.

Priznanie využitia generatívnej AI v akademickej práci

Použitie generatívnej AI musí byť potvrdené v časti „Čestné prehlásenie“, prípadne „Poďakovanie“ každej akademickej práce, kde bola použitá ako funkčný nástroj na pomoc pri procese vytvárania akademickej práce.

Minimálne požiadavky, ktorú treba zahrnúť do priznania je:

Názov a verzia použitého generatívneho systému AI: napr. ChatGPT-3.5.

Vydavateľ (spoločnosť, ktorá vytvorila systém AI); napr OpenAI, Google

URL systému AI.

Stručný popis (jednou vetou) kontextu, v ktorom bol nástroj použitý.

Napríklad:

Uvádzam na vedomie čitateľom mojej práce použitie ChatGPT 3.5 (Open AI, <https://chat.openai.com>) na zhrnutie mojich úvodných poznámok a na korektúru môjho konečného návrhu.

Ďalšie požiadavky môže stanoviť ústav, či fakulta, študijný program alebo jednotlivý pedagóg, a pre konkrétnu úlohu musia byť študentom jasné pravidlá priznania (rozsahu, spôsobu a pod.) pri zadávaní úloh. Ďalšie požiadavky môžu zahŕňať rozšírený popis v časti „Metódy“, ako napríklad:

¹ Poznámka: podľa poslednej konzultácie s OpenAI, ktorej obsah bol aktualizovaný naposledy ku septembru 2021, tento typ AI je stále vo vývoji a ešte neexistuje.

Ak je to relevantné, výzva použitá na vygenerovanie odpovede v systéme AI. Ak je to viac ako 3 oblasti pre použitie AI, tak navrhujeme uviesť rovnako ako „vyhľadávací algoritmus/stratégia“ pre vyhľadávanie zdrojov literatúry do prílohy práce celý prepis konkrétnych výziev zadaných AI pre vyhľadávanie vašich zdrojov alebo príkazy pre úpravu textu. Takto môže pedagóg jasne hodnotiť pridanú hodnotu AI a identifikovať aj vaše citovanie tohto zdroja (tak ako iných zdrojov v práci).

Dátum, kedy bol výstup vytvorený.

Získaný výstup (napr. „odkaz na chat“ v prípade ChatGPT alebo kompilácia všetkých výstupov vygenerovaných ako príloha).

Ako bol výstup zmenený na použitie alebo začlenenie do diela (napr. dokument so sledovanými zmenami alebo popisný odsek).

Tieto prílohy nemajú byť zahrnuté do počtu slov práce (teda rozsahu), pokiaľ nie je pre konkrétne zadanie alebo konkrétny študijný program alebo ústav/ fakultu stanovené inak. Poďakovanie, resp. čestné prehlásenie ku použitiu AI by malo byť umiestnené buď na začiatku alebo na konci dokumentu.

Citácie v texte a zahrnutie výstupu generovaného AI do zoznamu odkazov v samostatnej prílohe.

Niektoré štýly odkazovania naznačujú, že systémy AI by sa mali **citovať** podobným spôsobom ako iné zdroje, najmä osobná komunikácia, ale existujú problémy s citovaním systémov AI:

Nástroj AI nemožno klasifikovať ako autora, teda nemôže prevziať zodpovednosť za svoju prácu, ani nevytvára originálne nápady, ale reprodukuje nápady, ktoré sa našli inde a boli pripravené inými autormi.

Primárnou funkciou zoznamu referencií je, aby sa čitateľ mohol odvolávať na pôvodný zdroj, čo pri obsahu generovanom AI nie je možné.

Tento návrh preto uprednostňuje prístup väčšiny akademických vydavateľov, ktorý určuje, že systémy umelej inteligencie by sa nemali uvádzať ako autor, ani ako zdroj v zozname odkazov literatúry ale vo forme príloh (čím sa ani nezapočítava do rozsahu práce).

Môžu nastať prípady, keď je vhodné alebo potrebné, aby sa študent/ka alebo výskumník odvolával na výstup generovaný AI v rámci diela a/alebo ho zahrnul/a do referenčného zoznamu, napr. keď sa práca zaoberá témou generatívnej AI a diskusiou v súvislosti s výstupmi, kde je odkaz na formálne publikovaný výstup generovaný AI, kde to vyžaduje ústav/ fakulta alebo pedagóg, alebo kde študent/ka neidentifikoval primárny zdroj informácií napriek problémom so spoliehaním sa na generatívnu AI ako sekundárny zdroj informácií (čo možno považovať za zlú akademickú prax). Študentom sa odporúča, aby si to overili u svojho pedagóga resp. školiteľa.

V takýchto prípadoch by sa s výstupom malo zaobchádzať ako s dielom bez autora, generovaným AI, pokiaľ nie je uvedené inak v pokynoch pedagóga/ školiteľa alebo v štandardizovanom štýle odkazovania, ktorý používate.

¹ Poznámka: podľa poslednej konzultácie s OpenAI, ktorej obsah bol aktualizovaný naposledy ku septembru 2021, tento typ AI je stále vo vývoji a ešte neexistuje.

Používanie štandardizovaných štýlov odkazovania

Ak sa od študenta/ študentky alebo výskumníka/ výskumníčky vyžaduje, aby používal/a štandardný štýl odkazovania, ktorý má špecifické pravidlá pre citovanie a odkazovanie na AI, mali by sa okrem priznania používania AI, ako je uvedené vyššie, dodržiavať aj pravidlá štýlu citovania. To si môže vyžadovať použitie citácií v texte a položiek zoznamu referencií, kde je nástroj AI uvedený ako autor.

Príklad 1:

Keď sa zobrazí výzva: „Je delenie ľavého mozgu a pravého mozgu skutočné alebo metafora?“ text vygenerovaný ChatGPT naznačoval, že hoci sú tieto dve mozgové hemisféry do istej miery špecializované, „zápis, že ľudia môžu byť charakterizovaní ako predomancia pre „ľavý mozog“ alebo „pravý mozog“, sa považuje za prílišné zjednodušenie a populárny mýtus“ (OpenAI, 2023).

Odkaz k tomuto príkladu bude:

OpenAI. (2023). ChatGPT (verzia zo 14. marca) [Veľký jazykový model]. <https://chat.openai.com/chat>

Odkazy a citácie v texte pre ChatGPT sú **formátované nasledovne**:

OpenAI. (2023). ChatGPT (verzia zo 14. marca) [Veľký jazykový model]. <https://chat.openai.com/chat>

Zátvorková citácia: (OpenAI, 2023)

Naratívna citácia: OpenAI (2023)

Vždy si transparentne odkomunikujte so svojim školiteľom alebo pedagógom plán a použitie AI a jej priznanie vo svojej práci. Takýto prístup okrem toho, že je etický a akademicky prípustný, môže zvýšiť aj váš kredit v kombinácii prístupov pri písaní práce a inováciách. Vždy majte na zreteli potrebu kritického myslenia a overovania všetkých výstupov AI, tak s ohľadom na riziká ale aj poslednú aktualizáciu týchto nástrojov v roku 2021.

V Bratislave 21.2.2024

¹ Poznámka: podľa poslednej konzultácie s OpenAI, ktorej obsah bol aktualizovaný naposledy ku septembru 2021, tento typ AI je stále vo vývoji a ešte neexistuje.

Literárne zdroje:

Berkeley, I. S. N. (1997). What is Artificial Intelligence? The University of Louisiana at Lafayette.
<https://userweb.ucs.louisiana.edu/~isb9112/dept/phil341/wisai/WhatIsAI.html>

Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlström, P., Henke, N., & Trench, M. (2017). Artificial intelligence: The next digital frontier? McKinsey Global Institute.
<https://www.mckinsey.com/~media/McKinsey/Industries/Advanced Electronics/Our Insights/How artificial intelligence can deliver real value to companies/MGI-Artificial-Intelligence-Discussion-paper.ashx>

Cambridge University Press. (2017). Artificial intelligence. In Cambridge academic content dictionary.
<https://dictionary.cambridge.org/dictionary/english/artificial-intelligence>

Cambridge University Press. (2023). AI contributions to research content. In Authorship and contributorship.
<https://www.cambridge.org/core/services/authors/publishing-ethics/research-publishing-ethics-guidelines-for-journals/authorship-and-contributorship#:~:text=AI%20Contributions%20to%20Research%20Content&text=Any%20use%20of%20AI%20must,adequate%20citation%20and%20transparent%20referencing.>

Encyclopedia Britannica. (2023). Artificial intelligence. <https://www.britannica.com/technology/artificial-intelligence>

European Parliament. (2021, March 29). What is artificial intelligence and how is it used?
<https://www.europarl.europa.eu/news/en/headlines/society/20200827STO85804/what-is-artificial-intelligence-and-how-is-it-used>

High-Level Expert Group on AI. (2019). Ethics guidelines for trustworthy AI. European Commission.
<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

IBM. (2023). What is artificial intelligence (AI)? <https://www.ibm.com/topics/artificial-intelligence>

Intellectual Property Office. (2018). Copyright, Designs and Patents Act 1988.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/957583/Copyright-designs-and-patents-act-1988.pdf

Intellectual Property Office. (2021, January 4). Guidance. Exceptions to copyright.
<https://www.gov.uk/guidance/exceptions-to-copyright>

¹ Poznámka: podľa poslednej konzultácie s OpenAI, ktorej obsah bol aktualizovaný naposledy ku septembru 2021, tento typ AI je stále vo vývoji a ešte neexistuje.

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–
<https://doi.org/10.1016/j.bushor.2018.08.004>

McCarthy, J. (2007). What is artificial intelligence? <https://www-formal.stanford.edu/jmc/whatisai.pdf>

PWC. (2017). Bot.Me: A revolutionary partnership. How AI is pushing man and machine closer together.
<https://www.pwc.in/assets/pdfs/consulting/digital-enablement-advisory1/pwc-botme-booklet.pdf>

Ricketson, S., & Ginsburg, J. (2005). Appendix 1: Latest text of Berne Convention (1971 Paris Act plus Appendix). In *International copyright and neighbouring rights: The Berne Convention and beyond* (2nd ed.).
<https://global.oup.com/booksites/content/9780198259466/15550001>

Russell Group. (2023). Russell Group principles on the use of generative AI tools in education.
https://russellgroup.ac.uk/media/6137/rg_ai_principles-final.pdf

Turing, A. M. (1950) Computing machinery and intelligence. *Mind*, 49, 433–460.
<https://redirect.cs.umbc.edu/courses/471/papers/turing.pdf>

University of Cambridge. (2019, November 27). Plagiarism and academic misconduct.
<https://www.plagiarism.admin.cam.ac.uk/definition>

USCO. (2023, March 16). Copyright registration guidance: Works containing material generated by artificial intelligence. *Federal Register*, 88(51), 16190–16194. <https://www.govinfo.gov/content/pkg/FR-2023-03-16/pdf/2023-05321.pdf>

Vandeput, N. (2021, July 20). A brief history of neural networks. Medium. <https://medium.com/analytics-vidhya/a-brief-history-of-neural-networks-c234639a43f1>

World Intellectual Property Organisation. (1978). Guide to the Berne Convention for the protection of literary and artistic works (Paris Act, 1971).
https://www.wipo.int/edocs/pubdocs/en/copyright/615/wipo_pub_615.pdf

¹ Poznámka: podľa poslednej konzultácie s OpenAI, ktorej obsah bol aktualizovaný naposledy ku septembru 2021, tento typ AI je stále vo vývoji a ešte neexistuje.