

Smernice za uporabo UI – FF

(moška oblika se uporablja tako za ženski kot moški spol)

Preambula

Filozofska fakulteta Univerze v Ljubljani je oblikovala smernice za uporabo sistemov in orodij umetne inteligence (UI). Te smernice najprej opozarjajo na omejitve, nevarnosti in pasti uporabe teh orodij v znanosti, umetnosti, pedagoškem procesu ter tudi v družbi nasploh in s tem izpostavljajo bistveno vlogo preišljenosti pri uporabi UI. Hkrati so smernice namenjene spodbujanju tvorne, ustvarjalne, etične, varne in odgovorne uporabe sistemov umetne inteligence.

Nedavni razvojni skok velikih jezikovnih modelov je spodbudil zelo raznolike odzive na sisteme UI. Ti segajo od apokaliptičnih obetov in pozivov k ustavitvi ali upočasnitvi razvoja do ocen, da ne gre za nič novega ter da je vse, kar so ti sistemi zmožni, ponavljanje, prerazporejanje in nizanje besed. Vsekakor pa je ta skokoviti razvoj znova odprl vprašanje značilnosti, odlik in zmožnosti človekovega uma ter njegovih meja.

Raba umetne inteligence je priložnost za družbeni razvoj in razvoj znanosti, a hkrati odpira vprašanja etike akademskega dela, akademske integritete in standardov v znanosti in raziskovanju ter pedagoškem procesu. V akademskem prostoru se ne izogibamo uporabi novih orodij, do katerih privede znanstveni in tehnološki razvoj, zavzemamo pa se za njihovo odgovorno uporabo, ki bi vodila tako do izpopolnjenega življenja posameznikov kot tudi do napredka in blagostanja družbe kot celote. Zavzemamo se za to, da uporaba umetne inteligence ne bi privedla do poglobitve neenakosti in nesorazmerne porazdelitve bremen in koristi. Prav tako se zavedamo, da sistemi umetne inteligence ne nastajajo in ne delujejo ločeno od družbenoekonomskega in kulturnega okolja ter v tem oziru niso nevtralna tehnologija.

Z razvojem umetne inteligence in drugih novih tehnologij se soočamo s spremembami, ki neizogibno vplivajo tudi na način učenja, poučevanje in ocenjevanja. Generativna UI predstavlja nov izziv, saj s seboj prinaša tako priložnosti kot tveganja. Odgovorna in vodena uporaba teh orodij v smiselnem kontekstu lahko študentom pomaga pri učenju in razvijanju novih veščin, hkrati pa nenadzorovana raba generativne UI lahko privede do tega, da se zaobide ključne stopnje usvajanja veščin, pomembnih za doseganje učnih ciljev. Nadalje, razširjena uporaba in zanašanje na velike jezikovne modele lahko vplivata na strukture ustvarjanja znanja in avtoritete. Zato je pomembno, da izoblikujemo jasne smernice uporabe generativne UI v pedagoškem, raziskovalnem in strokovnem delu ter jim sledimo.

Splošna priporočila za rabo UI

Uporaba generativne UI je lahko zelo koristna, hkrati pa zaradi načina delovanja ter s tem povezanih možnosti napak predstavlja tudi tveganje. Veliki jezikovni modeli se res učijo na velikih količinah besedilnih podatkov, kar jim pomaga razviti razbiranje jezikovnih vzorcev in (v večini primerov) ustvarjanje kontekstualno ustreznih odzivov. Vendar pa nimajo možnosti neodvisnega preverjanja dejstev ali potrjevanja pravilnosti informacij, ki jih podajo. Pomembno je opozoriti tudi na to da veliki jezikovni modeli, ki lahko na prvi pogled ustvarijo smiselne in verjetne odgovore, seveda nimajo pravega razumevanja ali zavesti. Tak sistem UI ne razume vsebine, ki jo generira, hkrati pa ustvarja zelo tekoče besedilo, ki na prvi pogled deluje prepričljivo. To pomeni, da lahko:

- ustvari besedila, ki vsebujejo netočnosti, napačne ali zastarele informacije in zavajajoče vsebine,
- ustvari plagiat in krši avtorske pravice,
- navaja napačne ali neobstoječe reference,
- generira žaljive, neprimerne ali škodljive vsebine,
- generira pristranske vsebine,
- razkrije ali pravilno ugaane osebne informacije ali informacije, ki so občutljive zaradi drugih razlogov,
- se uporablja za namene manipulacije.

Generativna UI ne ločuje med tem, kaj je prav in kaj je narobe, saj nima zavesti in je amoralna. Zato je pomembno, da vse vsebine, ki jih ustvari generativna UI, preveri človek, in sicer nekdo, ki lahko kompetentno presoja o vsebini in primernosti njene uporabe.

Hkrati izpostavljamo, da razširjanje ustvarjenih vsebin velikih jezikovnih modelov sproža vprašanja o avtorstvu in izvirnosti človekovega dela in ustvarjanja. Pogosto je namreč težko razbrati, ali je vsebino ustvaril človek ali sistem umetne inteligence, kar lahko zmanjša vrednost in zaupanje, povezana s človeško ustvarjalnostjo in izražanjem.

Študijski proces je proces, ki študente usposobi za samostojno in kritično mišljenje ter opravljanje nalog brez pomoči UI. UI je tako **študentom in učiteljem** predvsem orodje za podporo temu procesu učenja in pridobivanju novih znanj in sposobnosti. V pomembni meri lahko UI olajša študij (npr. pregledovanje in izboljševanje besedil, računalniško programiranje, prevajanje, analize in interpretacije podatkov, podajanje razlag na različnih ravneh in na različne načine, priprava povzetkov in osnutkov besedil, generiranje povzetkov iskanj, generiranje grafičnih prikazov, slik, vizualizacije ...). Pri tem pa moramo biti pozorni, da tako pridobljene informacije niso nujno vedno pravilne oz. točne. Lahko so zavajajoče, izmišljene, stereotipne ipd. Zato je nujno, da je uporaba teh orodij transparentna in odgovorna. Za kakršno koli rabo UI in njenih tvorb je vedno odgovoren uporabnik.

Priporočila za učitelje

V času, ko narašča prisotnost UI na vseh področjih, je treba skrbeti za razvoj in spodbuditi različne pedagoške pristope, ki bodo temu prilagojeni, hkrati pa bodo upoštevali tudi predmetne specifične ter stopnjo študija. Pri sorodnih (ali istih) predmetih, ki jih uči več učiteljev, je nujno, da so pristopi čim bolj usklajeni. Pomembno je, da celoten proces morebitne rabe UI predvidimo, usmerjamo in nadzorujemo ter da s tem pričnemo že ob začetku izvajanja predmeta. Pri razvoju veščin dela z UI se je treba zavedati nekaterih ključnih omejitev generativne UI. Ta se izredno hitro spreminja in razvija, tako s stališča razvoja kot regulacije, zato se lahko zgodi, da bodo uporabljeni učni pristopi, metode ali celo veščine same do konca študijskega obdobja ali leta že zastareli. Učitelj, ki bo vključeval delo z UI, mora torej to področje skrbno spremljati in prilagajati svoje delo oz. usmerjanje študijskega procesa.

Učitelj naj nadzira in po potrebi usmerja uporabo UI pri študijskem procesu. Pomembno je, da se v okviru vsakega predmeta razvijajo tiste veščine in znanja, ki so predvideni v opisu predmeta ter so del samega študijskega programa. Pri tem se je treba zavedati, da lahko generativna UI zelo dobro izpolni ali reši nekatere vrste nalog. Kadar so te naloge ključne za razvoj omenjenih veščin in znanj, imamo na voljo

več možnosti. Kadar je to mogoče, lahko naloge preoblikujemo tako, da vsebujejo več kompleksnih korakov in dodatno spodbujajo kritično razmišljanje (na primer z dodatkom komentarja). Izvedbo nalog lahko zahtevamo tudi v živo ali v ustni obliki oz. v ustrezno kontroliranem okolju (npr. Exam.net).

Če sami uporabljamo sisteme UI, npr. v okviru ocenjevanja ali podajanja povratnih informacij, je pomembno zavedanje, da za vsako oceno in podano informacijo nosimo odgovornost sami ter moramo nujno bdeti na tem vidikom svojega dela.

V primeru suma, da je študent uporabljal nedovoljeno pomoč UI, npr. da je del besedila neposredni rezultat orodja UI, vendar to ni ustrezno navedeno ali utemeljeno, moramo od študenta zahtevati zagovor dela in dodatno dokazovanje avtorstva ter sprožimo ustrezne disciplinske postopke, če se naš sum potrdi.

Priporočila za študente

Študent se mora zavedati načina omejitev in načina delovanja (generativnih) orodij UI, hkrati pa se seznaniti s pravili rabe UI pri predmetu, ki so podana s strani učiteljev.

Mnoga od UI orodij ne povzemajo virov na način, kot smo ga vajeni pri akademskih besedilih, ki jih napiše človek. Tovrstna orodja se sicer učijo na obstoječih podatkih, a odgovore pripravljajo na podlagi statističnih jezikovnih modelov, ki temeljijo na pogostosti pojavljanja besed in besednih zvez. Zaradi tega lahko zlahka navedejo tudi napačne, zavajajoče ali nesmiselne informacije. Pri uporabi tovrstnih orodij je zato potrebna izjemna previdnost. Čeprav nekatera orodja povzemajo izvirne vire in jih navedejo na koncu odgovora, ni nobenih zagotovil, da so ti viri ustrezni ali ustrezno povzeti. Prav tako ti sistemi UI nimajo oblikovanega enotnega znanstvenega okvirja.

Pri pisanju besedil in opravljanju raziskovalnih nalog je potrebno vselej slediti načelom znanstvene etike in integritete. Študent je vedno dolžan prebrati vire, ki jih navaja, njihovo vsebino kritično ovrednotiti, jo ustrezno povzemati in navajati reference. Orodij UI zato ne sme nekritično uporabiti za pisanja daljših delov besedila. V skladu s študijskim redom morajo biti pisni izdelki, nastali v okviru študijskega programa, študentovo lastno avtorsko delo. Kot avtor strokovno-znanstvenega besedila je študent torej polno odgovoren za uporabo besedil, ki so generirana z orodji UI. Če je študent za pripravo besedila uporabil orodje UI v katerikoli obliki ali fazi dela, mora biti to jasno navedeno v besedilu ali opombah. Enako velja za predstavitve ter druge naloge.

Če učitelj eksplicitno ne odobri uporabe generativne umetne inteligence pri pouku, nalogah ali izpitih, uporaba teh orodij ni dovoljena. Če študent kljub temu uporablja tovrstna orodja, je to enako, kot če bi mu pri nalogah in izpitih pomagal nekdo drug. Učitelji bodo v primeru suma, da je strokovno-znanstveno besedilo, del besedila ali druga opravljena naloga neposredni rezultat orodja UI in to ni ustrezno navedeno v besedilu, od študenta zahtevali (ustni) zagovor dela, dodatne reference in dodatno dokazovanje avtorstva ter sprožili ustrezne disciplinske postopke.

Priporočila za raziskovalce

Pri uporabi orodij generativne UI oz. drugih metod strojnega učenja pri raziskovalnem delu je za ohranjanje akademske integritete nujno spoštovati enake etične standarde kot ob kateri koli drugi vrsti akademskega dela. To med drugim vključuje:

- ustrezno navajanje uporabljenih orodij, ki so ustvarjalno prispevala k vsebini dela, in priznavanje ustvarjalne vloge umetne inteligence v raziskovalnem in umetniškem procesu;
- preverjanje ter kritično ovrednotenje rezultatov, ki jih ustvarijo sistemi umetne inteligence, ter razumevanje njihove omejitve in morebitne pristranskosti;
- prevzemanje odgovornosti za vsebino dela in navajanje primarnih znanstvenih virov;
- odgovorno nalaganje podatkov, informacij in dokumentov v orodja umetne inteligence. Ne nalagajte osebnih podatkov, zaupnih dokumentov, dokumentov, za katere nimate ustreznih dovoljenj rabe, vključno z informiranim soglasjem, oziroma so zaščitene z avtorskimi pravicami itd.

Preglednost procesa in rezultatov dela v akademskem okolju, vključno z uporabo orodij, ki temeljijo na umetni inteligenci, je pomembna za ohranjanje zaupanja in verodostojnosti v delo akademske skupnosti. To zahteva ustrezno navajanje referenc pri uporabi besedil, pri katerih je uporaba umetne inteligence bistveno ustvarjalno prispevala k vsebini, kar pomeni jasno navedbo načina, obsega in namena uporabe UI pri raziskovalnem delu.

Priporočila za strokovne delavce

Pri delu strokovnih služb je uporaba orodij UI dobrodošla pomoč, saj lahko predstavlja dodaten vir informacij. Pri tem je pomembno, da se jih uporablja kot pomožna orodja. Dobljeni rezultati ne morejo predstavljati končne vsebine, temveč jih je treba vselej preveriti, pretehtati in kritično ovrednotiti. Etična in dopustna je uporaba orodij UI za jezikovne preglede ter podporo/pomoč pri branju in pisanju v tujih jezikih, za podporo pri iskanju idej, pri pripravi grobih osnutkov dopisov, kot pomoč za uporabo različnih računalniških programov, kot pomoč pri pripravi analiz ipd. Pri tem mora biti nalaganje podatkov, informacij in dokumentov v orodja UI skrajno odgovorno – ne nalaga oziroma ne vnaša se osebnih ali zaupnih podatkov, zaupnih dokumentov ali dokumentov, za katere nimamo ustreznih dovoljenj rabe oz. so zaščitena z avtorskimi pravicami.

Sklepni del

Z upoštevanjem zgoraj podanih izhodišč in načel lahko zaposleni in študenti na FFUL izkoristimo prednosti umetne inteligence ter hkrati ohranjamo akademsko integriteto in spodbujamo njeno odgovorno uporabo ter tako pozitivno vplivamo tudi na širšo skupnost. Za rabo orodij UI, vključno z vsemi njenimi izdelki, nosijo odgovornost njeni uporabniki.

V prihodnje gre pričakovati napredek v smeri še večjega obsega in učinkovitosti teh modelov, hkrati pa gredo težišča razvoja v smeri bolj pretanjenih nastavitvev in zmožnosti prilagajanja teh sistemov (glede na jezik, različna področja delovanja in modalitete, ciljne skupine uporabnikov, ipd.) ter vključevanje UI v druga, že obstoječega orodja in sisteme. Ta razvoj je potrebno skrbno spremljati, sproti uvajati nujne prilagoditve ter vseskozi tudi kritično vrednotiti njegove okvire in posledice.