

	NAVODILA	Oznaka:00S-Nav15-01 St. zaupnosti: PT- I Stran 1 od 6
	ZA UPORABO SISTEMOV GENERATIVNE UMETNE INTELIGENCE	

Razdelilnik: vsi zaposleni v NIB

Original hrani poslovni sekretar skrbnik dokumentov na ravni NIB

veljavna verzija dokumenta dostopna na strežniku na naslovu: P:\Dokumenti-NIB\Dokumenti-veljavni\NIB-00\Splošne-zadeve\Nav\00S-Nav15-01-Uporaba-generativne-umetne-inteligence

Obveščeni:

vsii zaposleni

Pripravili: dr. Damjana Miklič Milek, Katarina Kafadar dr. Damjana Miklič Milek Podpis: Katarina Kafadar Podpis: Datum: 21.10.2024	Velja od: 1.11.2024 Kopija natisnjena: 21.10.2024
Pregledala: Komisija za Etiko in enake možnosti Predsednica: dr. Andreja Ramšak Podpis: Datum: 21.10.2024	
Odobrila: direktorica NIB, prof. dr. Maja Ravnikar Podpis: Datum: 21.10.2024	Verzija izdaje: 01 Ime datoteke: Uporaba-generativne-umetne-inteligence

Na podlagi Živih smernic za odgovorno rabo generativne umetne inteligence v raziskavah, ki jih je marca 2024 sprejela Evropska komisija, Generalni direktorat za raziskave in inovacije (EU EC – DG Research and Innovation), na podlagi Strategije NIB, sprejete 27. 2. 2023 ter na podlagi Etičnega kodeksa NIB, sprejetega 22. 2. 2022, sprejme direktorica dr. Maja Ravnikar naslednja

NAVODILA

ZA UPORABO SISTEMOV GENERATIVNE UMETNE INTELIGENCE

na Nacionalnem inštitutu za biologijo

1. OPREDELITEV POJMOV

UMETNA INTELIGENCA¹ (ang. »Artificial intelligence«, v nadaljevanju **UI**) je področje računalništva in informatike, ki se osredotoča na razvoj sistemov, ki imajo sposobnost izvajanja nalog in običajno zahtevajo človeško inteligenco (npr. učenje ali sklepanje, tolmačenje in obdelava informacij, reševanje problemov, napovedovanje, odločanje, včasih pa tudi ustvarjanje idr.) (Kok idr., 2022²).

GENERATIVNA UMETNA INTELIGENCA (ang. »Generative artificial intelligence«, v nadaljevanju **GenUI**) so orodja, ki se uporabljajo za ustvarjanje novih vsebin, kot so npr. besedila, slike, zvoki, videoposnetki, programska koda idr. (Budhwar idr., 2023³). Primeri takih orodij za samodejno ustvarjanje besedil so npr. ChatGPT, Gemini, LLaMA, za ustvarjanje slik npr. DALL-E 2, DreamStudio, za ustvarjanje videoposnetkov npr. Synthesia ipd., na trgu pa se vsakodnevno pojavljajo nova orodja z novimi ali izboljšanimi funkcionalnostmi.

GenUI deluje tako, da se na podlagi ogromnih količin vhodnih podatkov, ki jih uporabnik poda algoritmu, nauči njihovih struktur in vzorcev, kar omogoča ustvarjanje novih (izhodnih) podatkov (besedil, slik, videoposnetkov itd.). Vhodne podatke GenUI pridobi z medmrežja (podatke s spletnih strani, družbenih omrežij itd.). Nove vsebine ustvari tako, da analizira npr. zaporedje in porazdelitev besed, slikovnih točk (pikslov) ali drugih elementov vhodnih podatkov. Rezultate generira na podlagi statističnih metod.

¹ [Nature portfolio](#) - Artificial Intelligence (AI).

² Kok, J. N., Boers, E. J. W., Kusters, W. A., Putten, P. van der in Poel, M. (2022). Artificial intelligence: definition, trends, techniques, and cases. V Knowledge for sustainable development: an insight into the Encyclopedia of life support systems (str. 1095–1107). Paris, London: UNESCO, EOLSS.

³ Budhwar, P., Chowdhury, S., Wood, G., Aguinis, H., Bamber, G. J., Beltran, J. R., Boselie, P., Cooke, F. L., Decker, S., Knoblich, A. J., Malik, A., Paauwe, J., Papagiannidis, S., Patel, C., Pereira, V., Ren, S., Rogelberg, S., Saunders, M. N. K., Tung, R. L. and Varma, A. (2023). Human resource management in the age of generative artificial intelligence: Perspectives and research directions on ChatGPT. Human Resource Management Journal, 33(3), 606–659. <https://doi.org/10.1111/1748-8583.12524>

2. NAMEN NAVODIL

Na NIB postavljamo v središče temeljna vprašanja etike raziskovalnega dela⁴, integritete in standardov v znanosti in raziskovanju. Vsekakor se v znanstvenem okolju ne smemo izogibati uporabi novih orodij, ki se pojavljajo ob tehnološkem razvoju, zavzemamo pa se za njihovo odgovorno uporabo, ki bo vodila v napredek celotne družbe.

Na NIB spodbujamo in se zavzemamo za odgovorno uporabo metod in orodij in sistemov UI.

GenUI lahko ustvari možnosti in tveganja, ki jih je danes težko predvideti in v celoti razumeti. Navodila NIB se bodo zato redno posodabljala, razvijala in prilagajala in tako ostala koristen vir informacij za odgovorno rabo.

Na NIB v nadaljevanju povzemamo »Žive smernice za odgovorno rabo generativne umetne inteligence v raziskavah«⁵ in se kot inštitut zavzemamo za odgovorno uporabo orodij GenUI.

3. AKTIVNOSTI NIB PRI ODGOVORNI UPORABI GenUI NA NIB

NIB se zavzema za odgovorno uporabo GenUI pri čemer:

1. SPODBUJA, USMERJA IN PODPIRA ODGOVORNO UPORABO GenUI V RAZISKOVALNI DEJAVNOSTI IN NAVODILA VKLJUČUJE V INTERNE PROCESSE DELA IN DRUGE INTERNE DOKUMENTE

- NIB zagotavlja in/ali omogoča usposabljanje o uporabi GenUI, zlasti (vendar ne izključno) o preverjanju rezultatov, ohranjanju zasebnosti, naslavljanju pristranskosti in zaščite pravic intelektualne lastnine.
- Vodstvo NIB na pobudo zaposlenih organizira posvete z zaposlenimi o uporabi GenUI.
- NIB navodila uporablja kadarkoli je to mogoče. Po potrebi navodila dopolni s posebnimi dodatnimi priporočili in/ali izjemami (npr. pravila EU o varstvu podatkov, zaščita intelektualne lastnine itd.).
- NIB zagotavlja ustrezno raven kibernetske varnosti sistemov, predvsem tistih, ki so povezani z internetom.

2. SPREMLJA UPORABO GenUI ZNOTRAJ SVOJE ORGANIZACIJE

- NIB se zaveda raziskovalnih dejavnosti in procesov, pri katerih se uporablja GenUI. Informacije o uporabi GenUI med zaposlenimi lahko NIB:
 - uporabi za zagotavljanje in posodabljanje nadaljnjih priporočil za uporabo GenUI,

⁴ All European Academies. [The European Code of Conduct For Research Integrity](#). Berlin.

⁵ https://research-and-innovation.ec.europa.eu/document/download/2b6cf7e5-36ac-41cb-aab5-0d32050143dc_en?filename=ec_rtd_ai-guidelines.pdf

- za prepoznavanje potreb po usposabljanju in razumevanju, kakšna vrsta podpore bi bila najbolj koristna za uporabnike GenUI na NIB
- uporabi za preprečevanje morebitne zlorabe orodij GenUI.

4. NAVODILA ZA RAZISKOVALCE

Za uporabo GenUI na odgovoren način, morajo raziskovalci:

1. PREVZETI ODGOVORNOST ZA ZNANSTVENE REZULTATE

- Raziskovalci so odgovorni za celovitost vsebine ustvarjene z orodji UI ali s pomočjo teh orodij. Vsebina ne sme biti izmišljena, ponarejena ali plagiat⁶.
- Raziskovalci ohranjajo kritičen pristop pri uporabi rezultatov, ki jih ustvari GenUI in se zavedajo omejitev orodij, kot so pristranskost, halucinacije in netočnosti⁷
- Sistemi UI niso niti avtorji niti soavtorji. Avtorstvo pomeni posredovanje in odgovornost, zato jo imajo človeški raziskovalci.
- Raziskovalci v znanstvenem procesu ne uporabljajo izmišljenega materiala, ki ga ustvari GenUI, na primer ne ponarejajo, spreminjajo ali manipulirajo z izvirnimi raziskovalnimi podatki.

2. UPORABO GenUI TRANSPARENTNO NAVAJATI

- Raziskovalci za namene preglednosti pri pisanju sistematskih preglednih člankov (iskanje literature, metodologije ipd.) podrobno navedejo, katera orodja GenUI so bila uporabljena v raziskovalnih procesih. Sklicevanje na orodje mora vključevati ime, različico, datum itd. ter kako je bilo orodje uporabljeno in kako je vplivalo na raziskovalni proces. Če je to primerno lahko raziskovalci dajo na voljo tudi vhodne podatke (namige) in rezultate v skladu z načeli odprte znanosti. V kolikor raziskovalci GenUI uporabljajo za slogovne in estetske popravke besedil, le-te ni potrebno navajati.
- Raziskovalci upoštevajo naključnostno naravo GenUI, ki je nagnjena k temu, da iz istega vhoda ustvarja različne rezultate. Raziskovalci si prizadevajo za ponovljivost in robustnost uporabljenih GenUI, vključno z morebitnimi pristranskostmi v ustvarjeni vsebini ter o možnih ukrepih za njihovo ublažitev.

3. ZAGOTAVLJATI ZASEBNOST, ZAUPNOST IN ŠČITITI PRAVICE INTELEKTUALNE LASTNINE

- Raziskovalci se zavedajo, da se lahko ustvarjeni ali naloženi vhodni podatki (besedilo, podatki, pozivi, slike itd.) uporabijo za druge namene na primer za usposabljanje modelov UI. Za zaščito neobjavljenega ali občutljivega dela (na primer svoje neobjavljeno delo ali neobjavljeno delo drugih) ni dovoljeno naložiti v spletni sistem

⁶ Evropski kodeks ravnanja za poštenost raziskav; <https://allea.org/wp-content/uploads/2023/06/European-Code-of-Conduct-Revised-Edition-2023.pdf>

⁷ generativni modeli UI hipoma generirajo koherentna, tekoča, človeškim podobna besedila. V hitro proizvedenih rezultatih pa se pogosto skrivajo napačna dejstva, polresnice, izmišljene reference in tudi popolne izmišljotine, ki jim pravimo "halucinacije". Na dnu začetne strani vmesnika ChatGPT se nahaja obvestilo, ki se nanaša na vse klepete s sistemom, da se lahko zgodi, da "ChatGPT generira netočne informacije o ljudeh, krajih ali dejstvih". Zanesljivost sistema ChatGPT je okoli 60 % ali slabša.

UI, razen če obstajajo zagotovila, da podatki, ne bodo ponovno uporabljeni npr. za usposabljanje prihodnjih modelov ali za neizsledljivo in nepreverljivo ponovno uporabo podatkov.

- Raziskovalci pazijo, da osebnih podatkov tretjih oseb ne posredujejo spletnim GenUI, razen, če jim je posameznik, na katerega se nanašajo osebni podatki dal soglasje in če imajo raziskovalci jasno določen cilj za katerega bodo uporabili osebne podatke, tako da so v skladu s pravili EU o varstvu podatkov ⁸
- Raziskovalci razumejo tehnične in etične posledice v zvezi z zasebnostjo, zaupnostjo in pravicami intelektualne lastnine. Preverijo na primer možnost zasebnosti orodij, kdo upravlja orodje (javne, zasebne ustanove, podjetja itd.), kje orodje deluje in kakšne so posledice za vse naložene informacije. To zajema zaprta okolja, gostovanja na infrastrukturi tretje osebe z zagotovljeno zasebnostjo, kot tudi odprte platforme, dostopne prek interneta.

4. SPOŠTOVATI VELJAVNO NACIONALNO ZAKONODAJO⁹, ZAKONODAJO EU^{10,11} IN MEDNARODNO ZAKONODAJO; zlasti rezultati, ki jih ustvari GenUI, so lahko še posebej občutljivi glede varstva pravic intelektualne lastnine in osebnih podatkov

- Raziskovalci so pri uporabi rezultatov GenUI pozorni na možnost plagiatorstva (besedila, slike, kode itd.). Raziskovalci spoštujejo avtorstvo drugih in po potrebi citirajo njihovo delo. Rezultati GenUI (na primer veliki jezikovni modeli) lahko temeljijo na rezultatih nekoga drugega, zato jih je potrebno ustrezno prepoznati in navesti.¹²
- Rezultati, ki jih ustvari GenUI, lahko vsebujejo osebne podatke. Če to postane očitno so raziskovalci odgovorni za odgovorno in ustrezno ravnanje z vsemi pridobljenimi osebnimi podatki pri čemer je treba upoštevati pravila EU o varstvu podatkov.

5. STALNO SPREMLJATI RAZVOJ UPORABE GenUI, z namenom čim bolj izkoristiti njihove prednosti

- Generativna orodja UI se hitro razvijajo in redno se pojavljajo novi načini njihove uporabe. Raziskovalci so na tekočem z najboljšimi praksami in jih delijo s kolegi, sodelavci in drugimi zainteresiranimi deležniki.

6. SE IZOGIBATI BISTEVNI UPORABI GenUI PRI OBČUTLJIVIH DEJAVNOSTIH^{13,14}, kar lahko vpliva na druge raziskovalce ali organizacije (na primer strokovni pregled besedil npr. člankov ali projektnih prijav ali njihovih delov, ocena raziskovalnih predlogov ipd.)

⁸ https://commission.europa.eu/law/law-topic/data-protection/eu-data-protection-rules_en

⁹ [Zakon o avtorski in sorodnih pravicah](#) (Uradni list RS, št. 16/07 – uradno prečiščeno besedilo, 68/08, 110/13, 56/15, 63/16 – ZKUASP, 59/19 in 130/22)

¹⁰ Evropska unija: [Etične smernice za uporabo umetne inteligence in podatkov](#). Bruselj

¹¹ European Research Council. [Current position of the ERC Scientific Council on Artificial Intelligence](#). Bruselj.

¹² Tehnologija se razvija in generativna orodja umetne inteligence bodo bolj sposobna zagotoviti pravilne citate. Orodja, ki zagotavljajo citate, bi lahko bila bolj praktična in bi jim lahko dali prednost, vendar je za navedbo in njeno pravilnost odgovoren raziskovalec.

¹³ Na primer, uporaba generativne umetne inteligence za iskanje informacij o ozadju za oceno ni bistvena uporaba, medtem ko prenos vrednotenja ali ocene prispevka je bistvena uporaba.

¹⁴ Notice Number NOT-OD-23-149. [The Use of Generative Artificial Intelligence Technologies is Prohibited for the NIH Peer Review Process](#). NIH Office of Extramural Research, Department of Health and Human Services (HHS).

- Izogibanje uporabi GenUI odpravi morebitna tveganja nepravilnega obravnavanja ali ocenjevanja, ki lahko nastanejo zaradi omejitev teh orodij (kot so halucinacije in pristranskost).
- Poleg tega bo to zaščitilo izvirno neobjavljeno delo kolegov raziskovalcev pred morebitno izpostavljenostjo ali vključitvijo v model UI (pod pogoji ki so podrobno navedeni zgoraj v navodilu za raziskovalce št. 3).

Končna določila:

NIB ocenjuje in pričakuje, da se bo področje GenUI hitro razvijalo, zato bo ta navodila redno pregledoval in jih po potrebi prilagajal in dopolnjeval.

direktorica NIB
prof. dr. Maja Ravnikar