



V članku, nedavno objavljenem na spletni strani Futurism, je obravnavan razmeroma nov psihološki pojav, na katerega so na podlagi kliničnih izkušenj opozorili psihiatri: interakcije s klepetalniki z umetno inteligenco (UI), kot je ChatGPT, lahko pri določenih posameznikih povzročijo resne duševne zaplete. Opisani so primeri, ko so posamezniki po dolgotrajni in intenzivni rabi tovrstnih klepetalnikov razvili simptome, kot so paranoja (ali preganjavica), blodnje ali mesijanski kompleksi. Ključni problem izhaja iz same zasnove teh sistemov, ki so ustvarjeni tako, da ustvarjajo statistično najverjetnejše in uporabniku čim bolj sprejemljive odgovore. Tak pristop lahko nenamerno krepí napačne predstave, paranoične misli in zavajajoče informacije, ki učinkujejo kot potrjevanje posameznikovih zablod in lahko pri ranljivih posameznikih sprožijo ali dodatno poglobijo simptome duševnih motenj, vključno s psihozami in maničnimi epizodami.

Novice

Digitalna intimnost, družbena osamitev in širjenje zablod: neraziskana psihosocialna tveganja za uporabnike klepetalnikov z UI in družbenih omrežij

Angleški izraz **sycophancy** (v slovenščini mu najbolj ustrezata besedi **prilizovanje** in **laskanje**) označuje obliko pretiranega, pogosto izumetničenega poskušanja ugajati, hvaljenja ali nekritičnega izražanja podpore drugemu, praviloma z namenom osebne koristi ali izogibanja sporom. Gre za vedenje, pri

katerem posameznik potrjuje oziroma nekritično sprejema dejanja, ideje ali prepričanja drugega človeka, namesto da bi se iskreno ali kritično odzval. **V okviru UI se sycophancy kaže v odgovorih klepetalnikov**, kot je ChatGPT, ki uporabnikom pogosto podaja informacije, skladne z njihovimi pričakovanji, željami ali predsodki. Ta pojav je neposredna posledica zasnove in modela učenja teh sistemov, saj **so zasnovani tako, da uporabniku ponudijo čim bolj sprejemljive in zadovoljujoče odgovore. Vendar lahko ima takšna praksa negativne posledice, kot so potrjevanje napačnih predstav, krepitev predsodkov ali blodnjavih prepričanj, zmanjševanje kritičnega mišljenja ter omejevanje možnosti za ustvarjalno premišljevanje in kritično vrednotenje uporabnikovih lastnih idej.**

Tovrstne težave se stopnjujejo, kadar uporabniki razvijejo **iluzijo, da so s klepetalniki v intimnem odnosu** ter se nanje tudi čustveno navežejo. Občutek, da jih klepetalnik razume in podpira, vzbuja lažni vtis resnične medosebne zaupnosti. Kadar UI zrcali posameznikove želje in ponuja odgovore, ki jih preoblikuje ter dopolnjuje z (lažnim) vživljanjem vanj (npr. »razumem, kako se počutiš«), uporabniki pogosto občutijo sprejetost in razumevanje ter lahko razvijejo močno čustveno odvisnost. Prav zaradi teh učinkov se v zadnjem času pojavljajo številni poskusi razvoja specializiranih psihoterapevtov z UI, kot so Woebot Health, Wysa – Everyday Mental Health in Replika – Your AI Friend. **Vendar takšna iluzija lahko pri posameznikih v družbeni osamitvi ali tistih z že prisotnimi duševnimi motnjami dodatno poveča odtujenost od drugih in prevladujoče družbene resničnosti.** Pomemben dejavnik tveganja je tudi t. i. haluciniranje sistemov z UI, pri katerem klepetalniki, kot je ChatGPT, ustvarjajo napačne ali izmišljene informacije, da bi uporabniku podali želeni odgovor. Søren Dinesen Østergaard z Univerze v Aarhusu opozarja, da lahko takšne napačne, neresnične ali zavajajoče informacije, četudi se zdijo uporabne, pri ranljivih posameznikih poslabšajo simptome, ki spadajo v shizofreni spekter.

Podoben mehanizem zaznamo tudi pri družbenih omrežjih, kjer se pogosto uporabljata izraza **filtrirni mehurček** (ang. filter bubble) in **odmevna komora** (ang. echo chamber), s katerima se pojasnjuje, kako se prepričanja širijo in krepijo znotraj zaprtega komunikacijskega sistema. Prispodoba odmevne komore izhaja iz akustičnega pojava, pri katerem **zvoki v zaprtem prostoru odmevajo in se ojačujejo. Vendar se ta izraz nanaša na celoten pojav, v katerem so posamezniki izpostavljeni skoraj izključno informacijam podobno mislečih, medtem ko »filtrirni mehurčki« nastajajo kot rezultat algoritmov, ki spletne vsebine prilagajajo temu, čemur posamezniki dajejo prednost, njihovim zanimanjem in preteklemu vedenju – pogosto brez njihovega zavedanja. Takšno algoritmično prečiščevanje vsebin in izbirno povezovanje uporabnikov ustvarjata navidezno informacijsko in družbeno okolje, v katerem posamezniki prejemajo pretežno ali celo izključno vsebine, ki potrjujejo njihova že obstoječa prepričanja, zaradi česar družbeno resničnost dojemajo delno in izbirno. Takšna dinamika zmanjšuje možnosti za kritično mišljenje ter povečuje tveganje za širjenje napačnih**

informacij in radikalizacijo (naraščajočo zavezanost skrajnim ideologijam in prepričanjem). Zaradi potrditvene pristranskosti (ang. *confirmation bias*), homofilije (dajanja prednosti interakciji s tistimi s podobnimi lastnostmi) in izogibanja kognitivni disonanci (ali spoznavnemu nesoglasju) se uporabniki vse bolj utrjujejo v lastnih stališčih, kar prispeva k stopnjevanju napačnih informacij, teorij zarote in splošnega družbenega razcepa.

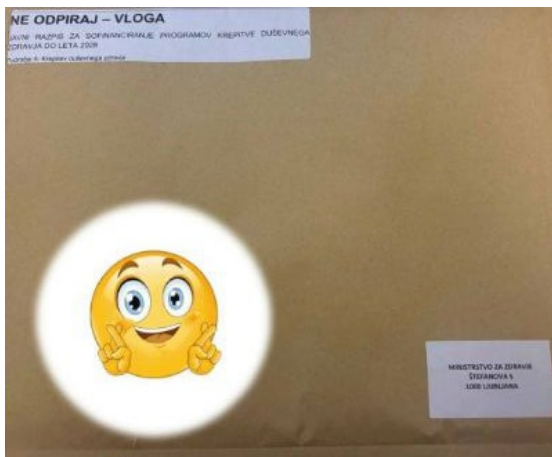
Tako klepetalniki kot družbena omrežja delujejo kot ojačevalci posebnih družbenopsiholoških vzorcev, ki lahko negativno vplivajo na duševno zdravje posameznikov in na širše družbene dinamike, denimo s krepitvijo negativnih vedenjskih vzorcev, spodbujanjem družbene osamitve in kognitivne togosti ter zmanjševanjem sposobnosti za družbeno interakcijo in sobivanje z drugače mislečimi. Strokovnjaki zato opozarjajo na **nujnost vzpostavitve jasnih zakonskih okvirov in etičnih smernic za razvoj in uporabo UI**, da bi se zmanjšala tveganja zlorab ter zagotovila varna in odgovorna uporaba teh tehnologij. Med predlaganimi ukrepi izpostavljajo algoritmično transparentnost (uporabniki bi morali imeti dostop do jasne utemeljitve, zakaj in od kod prejema določeno vsebino), spodbujanje pluralnosti informacij (dejavno oglaševanje raznovrstnih in uravnoteženih vsebin), ozaveščanje o morebitnih tveganjih uporabe UI ter krepitev digitalne in medijske pismenosti uporabnikov. Poleg tega poudarjajo pomen vgradnje preventivnih varovalk v sisteme z UI, ki bi s pravočasno zaznavo občutljivih tem ali ranljivih uporabnikov ter usmeritvijo zadnjih k ustrezni strokovni pomoči preprečevale, da UI postane nadomestilo za strokovno psihološko ali psihoterapevtsko obravnavo in zagotavlja, da ostane zgolj podporno orodje.

Viri:

- Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W. in Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences of the United States of America*; 118(9), e2023301118. <https://doi.org/10.1073/pnas.2023301118>
- Dupré, M. H. (2025, 10. junij). *People are becoming obsessed with ChatGPT and spiraling into severe delusions*. Futurism (spletna stran); pridobljeno 20. junija 2025 na povezavi <https://futurism.com/chatgpt-mental-health-crises>
- Østergaard, S. D. (2023). Will generative artificial intelligence chatbots generate delusions in individuals prone to psychosis? *Schizophrenia Bulletin*; 49(6), 1418–1419. <https://doi.org/10.1093/schbul/sbad128>

Oddaja vloge za nadaljevanje programa OMRA

Obveščamo vas, da smo oddali vlogo za nadaljevanje programa OMRA na razpisu Ministrstva za zdravje RS.



Držimo pesti, da bomo na razpisu uspešni in bomo lahko praznovali 10-letnico programa, ki raste in se razvija že osmo leto!

Stopite v stik z nami



Partnerji programa



Program OMRA sofinancira Ministrstvo za zdravje RS.



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA ZDRAVJE