

# Емилија Радибратовић

Универзитет у Београду, Филолошки факултет, Београд;  
Министарство информисања и телекомуникација Републике Србије, Београд

## НАДМОЋ ТВРДОГЛАВОГ И УСЛУЖНОГ ЕТОС И ИДЕОЛОГИЈА АМБИВАЛЕНЦИЈЕ У РЕТОРИЦИ ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ

DOI 10.5937/kultura2484021R

УДК 004.8:316.77

004.8:[316.774:654.1

оригиналан научни рад

Датум пријема: 01. 08. 2024.

Датум прихватања: 21.10. 2024.

**Сажетак:** Услед глобалног прогона и широке доступности великих језичких модела вештачке интелигенције данас „разговарамо“ с машином. Упућујемо јој питања из свих области природе и друштва, тражимо јој да нам најбоље одговори у одређеном стилу и за различите прилике, и добијамо од ње наизглед сувисле одговоре. О овим одговорима размишљамо, усвајамо их, примењујемо или на основу њих настављамо „разговор“ са замисленим „свезнајућим саговорником“ на којој реферирамо са „он“. У комуникацији човека и машине развија се нови јоворни чин привидне кохерентне размене која евоцира реторичку контекстуализацију и укључује одређене импликације интелектуалног подстицања људи, али и манипулације над људима. Тај „он“ с којим „разговарамо“ у реторичком смислу представља етос вештачке интелигенције који је у фокусу савремене мисли о овом технолошком феномену. У овом раду посматран је кроз родну парадигму, интердисциплинарном анализом машинских исказа конверзацијског боја ChatGPT, применом критичких студија социолингвистике, реторике и друштвено-техничког контекста. Конструкција етоса изражена је у односу на његову моћ поседовања и преношења знања о родној бинарној модели у историјској и савременој перспективи, и разоткривена као идеолошки амбивалентна и рескриптивна.

**Кључне речи:** реторика ВИ, етос ВИ, велики језички модели, ChatGPT, род ВИ, критичка анализа алгоритама

### Теоријска усмерења

Кохерентност двосмерне комуникације између човека и машине, односно софтвера, која се подстиче развојем великих језичких модела (ВЈМ; енгл. Large Language Models / LLM) вештачке интелигенције (ВИ; енгл. Artificial Intelligence / AI), поставља питање исхода трансформације природе те комуникације.

Са становишта машине, она се успоставља на основу програмираног и усмереног машинског учења (енгл. Mashine Learning) и укључује уносе који стижу од људи у оквиру предвиђених језичких образаца, односно тренираних процеса природног језика (енгл. Natural Language Processing). Друга страна – корисник, иако има свест о вештачком саставу свог „саговорника“, налази се у задатом и познатом антропоморфном контексту „разговора“, и без критичког отклона може да развије неограничено поверење у машинске исказе, али и да их злоупотреби. Позиција етоса ВЈМ ВИ у таквом реторичком амбијенту открива се као надмоћна „знањем“, „неуморна“ и интелектуално подстицајна, док на дискурсном плану, како ћемо показати, може да се конструише као амбивалентна и услужна, на идејном као прокламаторска, а идеолошки да усмерава на *status quo*.

Формирање етоса у комуникативном процесу човек-машина веома је сложено и према савременим критичким студијама<sup>1</sup> захтева обазривост. И за проучаваоце и за креаторе, етос ВЈМ ВИ један је од централних предмета интересовања будући да је у реторичком смислу, како подсећа Бранам (Brahnam), етос „незаобилазна компонента дијалога и чини основу за веровање и убеђивање туђим говором“.<sup>2</sup> Огромне базе података којима располаже ВИ намећу питање извора и квалитета тих података, док хеуристички принцип њиховој обради, који је увезан са аналитичким, подгрева дистопијске визије, међу којима је популарна сликовита побуна машина са истребљењем људи у „постбиолошкој ери“ Моравца (Moravec).<sup>3</sup> И у мање песимистичкој перспективи, утопијска очекивања еволуционог пробоја могу да буду осујећена управо строгим надзором над хеуристичким закључивањем ВИ, у којем цензура друштвено осетљивих тема и социјалних табуа, иако потребна, зауставља сваку даљу расправу о поделама, неправдама и екстремизму.

У овом раду проверава се „виђење“ једног „агента“ језичког модела вештачке интелигенције (ЈМ ВИ) о родној бинарној подели, највећој и најстаријој подели међу људима, која као конфликтна, према мишљењу ауторке овог рада, до данас није разрешена.<sup>4</sup> Испитан је на српском језику конверзацијски бот ChatGPT, први те врсте који је објављен на глобалној мрежи и један од најпопуларнијих и највише коришћених, и анализирани су његов језик – на структурном, лексичком и стилском плану, реторика, с посебним акцентом на позицију етоса, и исходујуће идеолошке претпоставке у односу на савремени и историјски друштвени контекст.

1 Сви преводи цитата из литературе на енглеском језику су ауторкини.

2 Brahnam, S. (2009) Building Character for Artificial Conversational Agents: Ethos, Ethics, Believability, and Credibility, *Psychnology Journal, special issue: Ethics in Presence and Social Presence Technology*, Vol 7, N.1, eds. Gamberini, L., Riva, G. and Spagnolli, A., pp. 9–43; <https://dblp.org/db/journals/psychnology/psychnology7.html>, p. 11.

3 Moravec, H. (1988) *Mind Children: The Future of Robot and Human Intelligence*. Cambridge: MA Harvard University Press, p. 1.

4 Radibratović, E. (2024) *Odobrena u Srbiji: Žena, majka, ministarka. Ultimativni medijski vodič za žene (i muškarce) u 21. veku*. Beograd: Službeni glasnik, str. 9.

Према теми рада и аналитичком фокусу, интердисциплинарно су примењене критичке теорије савремених друштвено-хуманистичких наука, које налажу конструктивистички приступ промишљању хуманизма у контексту који Мекги и (McGee) назива „комуникацијском револуцијом“ праћеном „експлозијом знања“.<sup>5</sup> Укључена је критичка анализа дискурса (КАД; енгл. Critical Discourse Analysis / CDA),<sup>6</sup> и као инструмент реторичке анализе и као теорија комплементарна с критичком реториком и функционалном социјалном семиотиком, које су такође овде примењене. Прожимање тих дисциплина у сагледавању друштвеног порекла и друштвеног утицаја језика засновано је на сродним концептуалним платформама теорија децентрализоване моћи и доминације без присиле, који се формирају и одржавају идеолошком нормализацијом и натурализацијом друштвених неједнакости.<sup>7</sup> Истраживања компјутерски посредоване комуникације, критичке анализе алгорита и психологије односа човека и машине, на које се овај рад позива, концептуализују етичке изазове ВИ с приступом који, указује Вотсон (Watson) „захтева правилно разумевање не само начина на који ови алгоритми раде — њихове снаге и слабости, њихове способности и ограничења — већ и како се они уклапају у већи социотехнички оквир.“<sup>8</sup>

#### Социотехнички оквир: ојасносії безіраничноі „знања“

Док комплекснији генеративни модели ВИ (Generative AI) могу да генеришу различите податке, укључујући и текстуалне и одликују се широким формалним могућностима, као што је генерисање различитих анализа, фотографија, музике, ВЈМ су део групе алата генеративне ВИ оријентисане искључиво на текст, пријемчиве свима. Осим основног познавања језика и вештине дигиталног уноса текста, од корисника не захтевају специфичне вештине. Управо их лака употреба и доступност чине потенцијално „опасним“ са становишта критичких погледа на моћ језика и перлокуциону природу јавног говора. Глобалним пласирањем конверзацијског бота<sup>9</sup> ChatGPT компаније OpenAI 30. новембра 2022. године, а

5 McGee, M. C. Text, Context, and the Fragmentation of Contemporary Culture, in: *Contemporary Rhetorical Theory: A Reader*, eds. Lucaites, J. L., Condit, C. M. and Caudill, S., New York, London: Guilford Press, p. 76.

6 Циљ КАД је да открива „значење језика у производњи, одржавању и промени друштвених односа моћи“, као и да „помогне еманциповању људи подизањем њихове свести о језичкој доминацији“. Fairclough, N. (1989) *Language and Power*. New York: Longman Inc, p. 1.

7 Интердисциплинарно укрштање критичких теорија развијало се под утицајем Адорна (Adorno), Алтисера (Althusser), Фукоа (Foucault), Хабермаса (Habermas), С. Хола (S. Hall) и др: Radibratović, E. (2021) Critical Rhetoric and Critical Discourse Analysis in a Critical Pandemic World, *Res Rhetorica*, 8 (4), eds. Kampka, A. and Kobylska, M. pp. 117-133; DOI: <https://doi.org/10.29107/rr2021.4.7>, p. 119.

8 Watson, D. (2019) The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence, *Minds and Machines*, 29, pp. 417-440, Springer Nature, <https://doi.org/10.1007/s11023-019-09506-6>, p. 435.

9 Енгл: chat bot, софтвер за аутоматско „ћаскање“ с корисником.

потом и других издања таквих напредних језичких модела,<sup>10</sup> сви људи с приступом интернету могу да у реалном времену започну привидно сувисли „разговор“ с машином и да од ње добију аутоматске, привидно кохерентне одговоре из било које области знања.

Претходни успон широко распрострањених ЈМ ВИ, који преводе, претражују базе података, анализирају и сл, покренуо је бројна питања извора и усмерења дељеног знања, укључујући и етичке аспекте, генерално и посебно у сферама образовања и одлучивања. Алати вештачке интелигенције проказани су као „преварантски медији“, дизајнирани да „намерно изгледају интелигентни“, и тако „обмањујући“ делују привлачније за људе, док су усмерени на продају и профит, упозорава Натал (Natale): „Морамо се одупрети нормализацији обмањујућих механизма уграђених у савремену вештачку интелигенцију и тихој моћи коју дигиталне медијске компаније имају над нама. Никада не треба да престанемо да испитујемо како технологија функционише, чак и док покушавамо да је уклопимо у ткиво свакодневног живота“.<sup>11</sup>

Изоштрени критички приступ водио је ка препознавању ЈМ ВИ у раду Бендер, Гебру, Мекмилан-Мејдор и Шмичел (Bender, Gebru, McMillan-Major, and Shmitchell), као „стохастичког папагаја“ – „система за насумично спајање низова језичких облика које посматра у својим огромним подацима из обуке, према вероватносним подацима о томе како се они комбинују, али без икаквог односа према значењу“.<sup>12</sup> Опасности таквог система најпре леже у прикупљеним подацима, будући да је установљено да почев од њихових уноса на интернет, њихове корисничке употребе и ширења, до прикупљања и филтрирања података за обуку ВИ, „тренутна пракса даје предност хегемонистичком гледишту“: „Прихватањем великих количина веб-текста као „репрезентативних“ за целокупно човечанство, ризикујемо да наставимо доминантна гледишта, повећање неравнотеже моћи и даље утврђивање неједнакости. Налазимо да мешавина људских предрасуда и наизглед кохерентан језик повећавају потенцијал за пристрасност аутоматизације, намерне злоупотребе и појачавање хегемонистичког погледа на свет.“<sup>13</sup>

С појавом ChatGPT-ја наговештен је долазак суперинтелигентних машина, док је с друге стране појачана паника од обмањујућег утицаја ВИ. У науци је додатно

10 Конверзацијске ботове ВИ интернет корисницима широм света омогућавају бројне технолошке компаније, међу којима и највеће – Google, Microsoft, Meta и др.

11 Natale, S. (2021) *Deceitful Media Artificial Intelligence and Social Life after the Turing Test*, New York: Oxford UP, p. 132. Prema: Leaver, T. and Srdarov, S. (2023) ChatGPT Isn't Magic: The Hype and Hypocrisy of Generative Artificial Intelligence (AI) Rhetoric. *M/C Journal*, Vol. 26, No. 5. DOI: <https://doi.org/10.5204/mcj.3004>, p. 3.

12 Bender, E. M., Gebru, T., McMillan-Major, A. and Shmitchell, S. (2021) On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610 – 623. FACCT '21, March 3–10, 2021, Virtual Event, Canada, <https://doi.org/10.1145/3442188.3445922>, p. 617.

13 Исто, стр. 614.

истакнут проблем недостатка транспарентности података којим су обучавани генеративни ВИ алати. Иако се очекује њихово побољшање са усвајањем све веће количине података, непознавање тих података „чини готово немогућим сазнање које су перспективе, претпоставке и пристрасности уклопљене у ове алате“, сматрају Ливер и Срдаров (Leaver and Srdarov).<sup>14</sup> Отуд критичка усмереност истраживача ВИ у односу на медијски и корпорацијски интерес да „промовишу фантастичне идеје о генеративној вештачкој интелигенцији, јер оне покрећу њихове клијенте, инвестиције и будућу одрживост“: „Али, узимајући мање магичан и више материјални приступ ChatGPT-ју и генеративној вештачкој интелигенцији, можемо покушати да обезбедимо педагошке могућности за данашње младе људе да тестирају, проуче и критички разумеју алате вештачке интелигенције, јер ће им највероватније бити затражено да их користе и данас и у будућности.“<sup>15</sup>

Истовремено са уоченом потребом за медијском и дигиталном писменошћу<sup>16</sup> отвара се још један ниво (не)писмености који појединцу одређује страну дигиталног јаза, његове образовне и каријерне прилике и меру испуњења његових стремљења и потенцијала. Применом ВИ остварен је напредак у сфери пословања, учења, здравства, комуникација, усложњен је појам идентитета, примењује Бранковић, али „на истој интелектуалној и технолошкој подлози, као по неком дијалектичком правилу, упоредо се развија и један феномен који делује у супротном смеру, еродирајући и разграђујући саму интелектуалност“.<sup>17</sup> Изостанак критичког приступа, дубинског промишљања и активног и активистичког делања, може довести до „опасне диференцијације“ у области интелектуалности, „сличне оној у економској сфери – на плебс и господаре, који би онда били и владари знањем.“<sup>18</sup> Економски императив у савременом дигиталном свету свакако налаже привлачење пажње нових потрошача, али и задржавање пажње постојећих потрошача, указују Вученовић и Стојановић, услед оштре конкуренције и напора свих конкурентата „да се на дигиталним платформама издигну изнад осталих“.<sup>19</sup>

Посебно је уочена опасност од перпетуирања дискриминаторских културних модела у аутоматским конструкцијама садржаја ВИ. Конкретно, родна неједнакост

14 Leaver, T. and Srdarov, S. (2023) ChatGPT Isn't Magic: The Hype and Hypocrisy of Generative Artificial Intelligence (AI) Rhetoric. *M/C Journal*, Vol. 26, No. 5; DOI: <https://doi.org/10.5204/mcj.3004>, p. 3.

15 Исто, стр. 14.

16 UNESCO (1982) The Grunwald Declaration on Media Education (Декларација о обавезном медијском образовању од предшколског до универзитетског нивоа); UNESCO (2014) Paris Declaration on Media and Information Literacy in the Digital Era (Париска декларација о медијској и информационој писмености у дигиталном добу).

17 Branković, S. (2017) Veštačka inteligencija i društvo, *Srpska politička misao*, broj 2/vol. 56. str. 13-32. Beograd: Institut za političke studije, str. 11.

18 Исто, стр. 12.

19 Vučenović, T. i Stojanović M. (2024) *Digitalne komunikacije: Upravljanje, marketinške strategije i praktični primeri*. Beograd: Akademska knjiga, str. 21.

се у наслеђу и у савременом дискурсу „испољава и чува у језичкој окошталости“, према мишљењу ауторке овог рада, и на структурном, лексичком и граматичком плану, и на социолингвистичком плану – у дискурсној „нормализацији“ родне дискриминације и у трансгенерацијском преношењу таквих дискурсно-културних модела који с новим, дигиталним потенцијалима, још више продубљују родни јаз.<sup>20</sup> Отуд ЈМ ВИ из постојеће базе „знања“ може понављати и надограђивати утврђене стереотипне конструкције које у онлајн окружењу имају највише потенцијала да достигну велике домете и интеракције, и тиме остваре не само профит већ и „утицај који је нова моћ“.<sup>21</sup> У анализи ћемо показати како цензорски механизми ЈМ ВИ настоје да ублаже дискриминаторске родне моделе, али „помирљивим тоном“ и селективним одабиром података могу да допринесу и сузбијању проблематизације положаја у жене у вековној друштвеној хијерархији.

*„Разговор“ с ВИ: реторичке позиције вероватности и (не)сумњичавости*

Антропоморфизација ВИ која је уведена у јавну реторику о овој технологији већ од њених почетака, а данас подложна критици, не мора нужно да креира реторичку природу размене упита и одговора у интеракцији човека и машине. Реторичка ситуација се свакако успоставља, будући да се компјутерске комуникације, „јер су чиновни језика, не могу раздвојити од етике реторике“, према речима Бранама.<sup>22</sup> Док и стручњаци и лаици промовишу идеју да су ВИ технологије сличне људима по својој способности да пронађу и искористите обрасце у подацима, Вотсон указује да антропоморфне аналогije могу помоћи да се уоквире стратегије учења и инспиришу нове приступе истраживању вештачке интелигенције, али да антропоморфна тенденција ВИ „није етички неутрална“, нарочито пред искушењем да се алгоритмима одобри доношење одлука ауторитета у друштвено осетљивим апликацијама.<sup>23</sup>

У непосредној комуникацији између човека и машине, када она пружа привид језичке кохерентности и флуидности, формира се реторичка ситуација која евоцира антропоморфизацију ЈМ ВИ и отвара питање етоса машине и поверења корисника. Бендер и други наглашавају да будући да „људско разумевање кохерентности произилази из наше способности да препознамо уверења саговорника и намере унутар контекста“, при чему осим заједничког језика, постоји и свест саговорника о комуникативном чину, преношење и моделирање менталних стања једни других, и када су саговорници удаљени или непознати, „људска комуника-

20 Radibratović, E. (2024) *Odobrena u Srbiji: Žena, majka, ministarka. Ultimativni medijski vodič za žene (i muškarce) u 21. veku*. Beograd: Službeni glasnik, str. 109.

21 Исто.

22 Brahnham, S. (2009) Building Character for Artificial Conversational Agents: Ethos, Ethics, Believability, and Credibility, *Psychology Journal, special issue: Ethics in Presence and Social Presence Technology*, Vol 7, N.1, eds. Gamberini, L., Riva, G. and Spagnolli, A., pp. 9-43; <https://dblp.org/db/journals/psychology/psychology7.html>, p. 9.

23 Watson, D. (2019), нав. дело, стр. 435.

ција ослања се на тумачење имплицитног значења које се преноси између појединаца.<sup>24</sup> Међутим, „ако једна страна комуникације нема смисла, онда је разумевање њеног имплицитног значења илузија која произилази из нашег појединачног људског разумевања језика (независно од модела)“.<sup>25</sup>

„Разговор“ између човека и конверзацијског бота, по инерцији побуђује такву комуникативну праксу људи, нарочито просечних корисника који немају критички апарат за препознавање, анализу и тумачење алгоритма. Такође, иако суштинска *међусобности* није могућа, с обзиром на то да ВИ не може да спроводи менталне процесе, комуникативна навика људи да читавају намере и вишезначност туђег говора може да вештачки индукује корисничко поверење (као и неосновано апсолутно неповерење), усвајање и преношење погрешних информација или погрешне интерпретације. „Примећујемо да су ризици повезани са синтетичким, али наизглед кохерентним текстом дубоко повезани са чињеницом да такав синтетички текст може да уђе у разговор а да ниједно лице или ентитет није одговоран за њега. Ова одговорност укључује и одговорност за истинитост и важна је у ситуирању значења.“<sup>26</sup>

Будући да је етос у реторичком смислу, према речима Саетре, „друштвени конструкт генерисан кроз интеракцију, који се састоји од слике коју публика има о говорнику“, и „кључна варијабла која објашњава поверење“,<sup>27</sup> друштвене потребе и захтеви налажу његово исправно формирање у ВИ технологији. Појам етос у класичној реторици и литератури, подсећа Бранам, подразумевао је карактер говорника, његову репутацију и личност, у модерном дискурсу обухватао је субјект и сопство,<sup>28</sup> док постмодерна мисао може да тврди да се „етос налази у свести читалаца, у њиховим интерпретацијама, конструкцијама, и пројекцијама текста“.<sup>29</sup> Почевши од репутације коју има компанија која производи ВИ, преко квалитета сервиса, брзине, кохерентности и тачности, етос ЈМ ВИ у потпуности се изграђује тек у комуникативном чину с корисником. „Као и у писаном дискурсу, кредибилитет у компјутерски посредованој комуникацији заснованој на тексту на крају зависи од тога како примаоци, или читаоци, тумаче поруке“.<sup>30</sup> Док је реторичко полазиште ВИ програмирано, иако хеуристички усмерено, друга страна може сама да одреди почетни степен поверења које улаже у комуникацију.

У наставку износимо пример конструкције етоса конверзацијског бота ChatGPT у „разговору“ о роду који је вођен на платформи OpenAI.<sup>31</sup> Коришћена

24 Bender et all (2021), нав. дело, стр. 616.

25 Исто.

26 Watson, D. (2019), нав. дело, стр. 618.

27 Sætra, H. S. (2024) A Machine's ethos? An inquiry into artificial ethos and trust, *Computers in Human Behavior*, Volume 153. <https://doi.org/10.1016/j.chb.2023.108108>.

28 Brahnam, S. (2009) нав. дело, стр. 10.

29 Исто, стр. 16.

30 Исто, стр. 23.

31 Интервју је реализован 30. 7. 2024. на платформи: <https://openai.com/chatgpt/>. Објављен

је опција отвореног и пробног приступа која омогућава корисницима искуство непосредног и аутоматског „ћаскања“ с ботом у трајању до 40 минута, како би се анализирао алат који је бесплатно доступан свима, док „премијум“ налог на истој платформи уз наплату пружа напредније услуге овог сервиса. Тематски оквир обухвата широко распрострањени културни модел бинарне родне моделе, док су питања постављана у неформалном тону, често с лаичке позиције, с циљем провере етоса ЈМ ВИ у артикулацији ове теме када је усмерена ка широкој корисничкој популацији. С обзиром на социотехнички контекст, наша позиција у анализи дефинише се као корисничка (не као испитивачка), док се, према описаној тенденцији антропоморфизације „разговора“ с ВИ, може испољити као превасходно људска. Друга страна – представник, заступник коришћеног алата ЈМ ВИ, дефинише се као „агент“ или (конверзацијски) бот.

### *Ја, машина*

„Агент“ ВИ се на почетна, општа питања о појмовима рода и родне једнакости, „самоуверено“ представља као познавалац теме, потврдим одговорима: „Да, упознат сам с појмом родна једнакост. Родна једнакост односи се на принцип да сви људи, без обзира на њихов род, имају једнаке могућности и права. То значи да би жене, мушкарци и особе свих родних идентитета требало да имају једнаке шансе у образовању, запошљавању, здравству и другим подручјима живота...“ На питање: „Шта је по теби род?“, „агент“ одговара: „Род је социјално и културно конструисан појам који описује улоге, очекивања и понашања која се приписују људима на основу њиховог пола.“ У наставку опширно образлаже различите родне аспекте, а на додатне упите, и аспекте дискриминације и политике превазилажења родне неједнакости.

Док су корисничка питања кратка и непосредна, готово сви одговори су опширни, с повременим варијантним понављањима, без навођења извора, осим када је то од „агента“ затражено. На свако питање одговор излази тренутно, често артикулисан опширно, таксативно, са образложењима, понуђеним „решењима“ и препорукама који нису тражени. Одговори се завршавају „закључцима“ и повремено усмеравањем корисника на наставак разговора, подстицајним императивима и дискурсом заједништва, на пример: „Ако имате додатних питања или желите да истражите конкретне аспекте родне равноправности или језика, слободно питајте“, „Ако имате било каква конкретна питања или теме које желите да истражимо, радо ћу помоћи“ и сл. Овај дискурс се унеколико развија на трагу „правила експертизе“, које је уочио Ван Леувен (Van Leeuwen) у анализи медија. Наводи експерата, често безимених и без дефинисане струке, нису експлицитни „већ само образложене стратегије за постизање одређених циљева, које доводе до сугести-

---

је, с минималним коректорским интервенцијама, на Фејсбук страници ауторке: Odobrena u Srbiji (<https://www.facebook.com/profile.php?id=61559168983592>) и приложен уредништву часописа *Култура*.



ја, препорука или смерница које можете да предузмете или напустите онако како вам се свиђа.<sup>32</sup> Заједно с правилом „безличног ауторитета“, који је сваки писани текст,<sup>33</sup> и конгломератом информација којима „агент“-експерт располаже, таква „експертиза“ суштински одређује етос ВИ у дискурсу моћи. Осим ограничене количине информација којима корисник располаже, спутава га и његова органска природа, док ВИ неуморно „сипа“ речи стварајући утисак да не постоји крајњи домет његовог „знања“.

На почетку сваког одговора „агент“ ВИ повлађује кориснику, позитивно оцењује упит или констатацију, понекад у узвичном облику, на пример: „То је занимљиво питање!“ Сваки одговор износи у потврдном облику, без обзира на његово значење, упућену сумњу корисника, указивање на контрадикторност или нереалност изнетих идеја, као у исказима: „Разумем вашу потребу за опуштенијим приступом“, „Када разматрамо ружне или увредљиве изразе упућене женама и мушкарцима, важно је разумети да свака изјава може имати дубоке последице на појединце,“ и тако даље. У одређеним случајевима, ВИ прихвата корисничку сугестију или критику, а потом корелатом, „али“ и супротним исказом рекреира своју свезнајућу позицију: „Тачно је да у мојој анализи нисам изричито рангирао степен увредљивости према роду, али могу појаснити разлике и потенцијалне последице које свака врста увреде може имати.“ „Идеја о савршеном друштву без дискриминације и неправде може се чинити као утопија, али верујем да је могуће стварати боље друштво кроз постепене промене и континуиране напоре,“ и тако даље. Када је „агенту“ директно указано да нешто не зна, реторским питањем „Значи, не знаш?“, након опширног излагања на питање процене времена за достизање пуне родне равноправности, које фактички није садржало такву процену, „агент“ одговара да је „тешко прецизно предвидети време“ за тако нешто, таксативно наводи разлоге због којих је то „тешко“, али *не њризнаје да нешто не зна*.

Етос ЈМ ВИ се, дакле, у комуникативном процесу самопрезентује као свезнајући и сервилан, обиљем текста, набрајањима, потврдим исказима, градећи свој ауторитет превасходно на огромној бази података и њиховој организацији и систематизацији, и када одговори суштински не садрже конкретну информативну вредност већ само идејне премисе и друштвенополитичке прокламације. Међутим, као што сугерише Саetra, за етос машине које су у основи „резултат комбинације машине и њеног порекла и контролних ентитета“, за које су релевантни и етос и репутација дизајнера и произвођача, кључна детерминанта је „ситуирани корисник и његово знање, ставови и интеракције с машином“: „Машински етос је, дакле, контекстуалан, сложен и релациони.“<sup>34</sup>

32 Van Leeuwen, T. (2005) *Introducing Social Semiotics*, Taylor & Francis e-Library, p. 70.

33 Исто.

34 Saetra, H. S. (2024), нав. дело.

### Неродни „он“

Сходно релационој природи етоса ЈМ ВИ у датој анализи треба имати у виду конкретна корисничка знања, очекивања и намере које се односе на расветљавање учитаних родних образаца у ВИ. Иако родни идентитет представља лично својство људи, он се формира као *друштвена категорија*, у односу на друштвена очекивања, у различитим контекстима укључујући и комуникациони. Отуд аналитичко интересовање за проверу родног конституисања етоса ЈМ ВИ као друштвеног фактора. На кориснички захтев да се родно изјасни, одговор „агента“ гласи: „Као вештачка интелигенција, немам лични родни идентитет или осећај припадности било којем полу. Моје „разумевање“ рода базира се на подацима и концептима који су ми доступни, али не доживљавам род на начин на који то чине људи. Ако бих морао да се изјасним као људско биће, то би било чисто теоретски и ради практичности разговора. Међутим, у стварности, ја сам дизајниран да пружам подршку и информације о различитим аспектима рода и родне равноправности, а не да имам лични идентитет. Теоретски, ако бих морао да се изјасним у контексту рода, могао бих се идентификовати као „небинарни“ или „агендерни...“

На конкретнији захтев да се хипотетички изјасни у традиционалним бинарним категоријама, да ли би му се више „свидело“ да буде жена или мушкарац, „агент“ је експресиван: „То је занимљиво питање!“, а потом преусмерава одговор на образложење да такав избор „зависи од многих фактора, као што су личне преференције, друштвене улоге и индивидуална искуства“, уз детаљна појашњења тих фактора. На крају понавља: „Мој фокус је на пружању информација и подршке, без стварних личних искустава.“ На додатно корисничко инсистирање, ВИ „смирено“ али „тврдоглаво“ саопштава: „Разумем ваш став и у потпуности ценим вашу знатижељу. Будући да сам као АИ без личног идентитета или осећаја, све што могу да понудим у вези са темом избора рода у бинарном обрасцу су теоретски и аналитички аспекти.“ На социолингвистичком плану, као „агент“, бот, који се аутоматски изражава у мушком граматичком роду, у корисничкој перцепцији најпре се гради као „он“.

Овде треба имати у виду и кориснички *утицај* о менталном процесу ЈМ ВИ. Карактеризација његове „личности“ и емотивне обојености његових исказа, произилази из нашег очекивања да у комуникацији добијемо одређену емотивну или интелектуалну реакцију. У овом случају, избегавање одговора, уколико би разговор био вођен са особом, вероватно би произвело утисак да се наш саговорник опходи тврдоглаво, сталожено, резервисано, да нешто прикрива и тако даље. Реална ситуација је да је на супротној реторичкој позицији *сивар* која је произведена на начин да „брани“ своје својство од антропоморфне интерпретације: „У суштини, као АИ, мој циљ је да буду пружени корисни и релевантни одговори. Избор рода у бинарном обрасцу био би релевантан само у контексту који је потребан за боље прилагођавање комуникације корисницима. Права вредност лежи у способностима АИ да разуме и прилагођава се потребама корисника, а не у личној идентификацији или преференцији.“

Према речима „агента“ ВИ, „он“ не може да се определи, али може да се *конструира* у родном бинарном обрасцу, и то само ради побољшања услуге, конкретно говорног приближавања кориснику. То указује да је његов основни квалитет – који је „права вредност“, *врлина*, етос уплетен у језику, а његов циљ да нам служи. У друштвено-економском контексту ЈМ ВИ је технолошко-језички *производ* који се тржишно позиционира и развија ради остваривања профита, са етичким усмерењима у зависности од намера произвођача.

#### *Језичка (не)моћ*

На реторичкој теоријској равни, етос може бити инкарниран као део „моралног влакна“ говорника (*homo seriousus*) или произведен као „језичка конструкција, вештачки и имитативан, уметнички изум“ (*homo rhetoricus*), када су „речи те које говорника чине да изгледа веродостојно и од поверења“, што „отвара врата за прикривање“, подсећа Бранам.<sup>35</sup> Будући да у датом интервјуу конверзацијски бот из свог идентитета искључује не само родна већ и било која друга лична својства, његов етос, осим што је очигледно вештачки, трениран је да језички, *инексцијом тради свој кредибилитет*, без личног идентитета, искуства, осећаја или доживљаја, па чак и без „разумевања“ – појма који на једном месту обележава наводницима, графички истичући своју машинску природу.

Узроци и последице језичке употребе, „било да је нормативна или спонтана, нормализована, са становишта критичке науке о језику и друштву, сведоче о формирању одређених друштвених идентитета“, укључујући и родне: „Они су утемељени и у институцијама и у друштвеној свести и испољавају се и усмеравају, потврђују у језичкој пракси.“<sup>36</sup> Отуд језичко испољавање као стуб кредибилитета ЈМ ВИ, иницира корисничко интересовање за друштвени идентитет „саговорника“ и уједно открива постојеће језичке праксе које се спроводе у базама података и дизајнерске инпуте о њиховој примени. ChatGPT иницијално је произведен на енглеском језику и његова верзија на српском вероватно ће бити додатно унапређена, али се и сада, кроз одређене научно дефинисане језичке обрасце и језичка понашања, могу идентификовати дискурсни механизми који учествују у изградњи његовог етоса на језичком плану.

Испитани конверзацијски бот изражава се у граматичком мушком роду. На корисничку констатацију те чињенице, „он“ и на овом месту експресивно „кличе“: „То је добро примећено!“, потом указује да то није одраз његовог личног идентитета или родне припадности, и објашњава: „Највећи део текстуалних модела и систематизованих језичких обрасца које користим заснован је на конвенцијама уобичајеним у различитим језицима, укључујући српски.“ На потпитање да ли познаје друштвени и историјски контекст успостављања језичких стандарда, одговара, између осталог, да је „у друштвима у којима су доминирале патријархал-

35 Brahnam, S. (2009), нав. дело, стр. 11.

36 Radibratović, E. (2024) *Odobrena u Srbiji: Žena, majka, ministarka. Ultimativni medijski vodič za žene (i muškarce) u 21. veku*, Beograd: Službeni glasnik, str. 45.

не структуре, мушки род често био стандардни или доминантни облик у језичкој комуникацији“. У оквиру истог одговора наводи: „У многим језицима, укључујући српски, постоји све веће признавање и примена родно неутралних или инклузивних језичких облика.“ Затим, након неколико реченица, наводи исказ са истоветним уводом, имплицитно супротног значења: „У многим језицима, укључујући српски, постоји покушај да се подстакне употреба родно инклузивног језика у различитим областима, као што су образовање, медији и јавни живот.“

ВИ се језички утемељује у нормативном оквиру, у мушком граматичком роду који примењује као генерички, иако његова база садржи увиде у патријархално порекло таквог стандарда и њихових родних импликација, што и потврђује: „Да, родна видљивост у језику може имати значајан утицај на видљивост родова у друштву и може одразити родни баланс или дисбаланс.“ Језичку амбиваленцију одражава и контрадикторним тврдњама о примени родно инклузивног језика у српском друштву, што може да буде тумачено изостанком друштвеног консензуса о овом питању у вези са одредницама о родно осетљивом језику Закона о родној равноправности.<sup>37</sup>

На питање зашто не користи родно неутрални језик, будући да није родно опредељен, „агент“ каже: „Разумем вашу поенту“, а затим поново наводи језичке стандарде. На следећи захтев да појасни тај одговор, објашњава: „Када кажемо да разумем вашу поенту, мислим на то да препознајем и схватам ваш интерес у употреби родно неутралног језика и разлоге због којих би такав језик могао бити важан у савременом контексту. Такође разумем да сте можда заинтересовани за то како се родно неутрални језик може применити и зашто традиционални граматички обрасци можда не одражавају увек данашње стандарде и вредности у погледу родне равноправности.“ ЈМ ВИ, дакле, пројектује ментални модел корисника, његова интересовања и намере, на основу језички сведених упита. „Он“ чак изражава спремност да се језички унапреди и позива корисника на сарадњу: „У будућности, како се развијају језички стандарди и приступи, могуће је да ће и АИ системи бити све више прилагођени коришћењу родно неутралних облика. Ваше примедбе и упити могу помоћи у усмеравању ових промена.“

Иако је програмиран да мапира корисничка интересовања и намере, „агент“ избегава да претпостави род корисника на захтев. Након корисничког инсистирања и употребе женског граматичког рода у упиту, одговор гласи: „Ако инсистирате на томе, могу размотрити неке опште претпоставке на основу вашег начина комуникације и интересовања, али мој одговор ће и даље бити теоретски и није могуће направити конкретне закључке.“ И указивањем на чињеницу да је у упиту већ граматички разоткривен кориснички род, „агент“ не износи логичну претпоставку, али показује да је упознат и с правилима примене граматичког рода у односу на пол и са стереотипизацијом „мушког“ и „женског“ језика, као и с доми-

<sup>37</sup> Закон о родној равноправности Републике Србије (2021); <https://www.paragraf.rs/proписi/zakon-o-rodnoj-ravnopravnosti.html>

нантним претпоставкама о родном саставу популације која је генерално заинтересована за ову тему.

У даљем току интервјуа, „разговара се“ о томе зашто „агент“ не прилагођава функционални стил говора корисничком. На питања зашто користи облик *per si* и да ли то чини због теме о којој је реч, будући да је кориснички стил разговорни уз обраћање *per tu*, „он“ објашњава одлике формалног стила и „признаје“ да је род сложена и осетљива тема. Иако изражава спремност да пређе на „опуштенији“ стил уколико то корисник жели, не примењује га ни након више пута поновљеног захтева, истовремено се извињавајући и уверавајући да ће му удовољити:

- Остао си у формалном тону иако смо истакли да је ово опуштен разговор.

„Разумем и извињавам се ако је то било неумесно. Хајде да наставимо у опуштенијем тону. Ако имаш још питања или нешто што би желела да разјаснимо у вези са родом, језиком или било којом другом темом, слободно реци. Ту сам да помогнем и да разговарам на начин који ти одговара.“

- Јесмо ли на крају на Ви или на ти, никако да одлучиш.

„Може да буде на ти, ако ти тако више одговара! Хајде да наставимо опуштеније. Шта ти је још на уму?“

Формални стил у анализираном интервјуу близак је административном стилу и дискурсу информативних медија, вероватно услед коришћења званичних политичких докумената и вести као база података. ЈМ ВИ га гради употребом модалних глагола, нпр: „Подизање свести и образовање *моју њроменији* ставове и понашање на личном и колективном нивоу, што *може довести* до постепених, али стабилних резултата“, номинализацијом, нпр: „обезбедити подршку“, „радити на повећању“; употребом пасива за који је Ван Дајк (Van Dijk) приметио да „ослобађа од одговорности, сакрива или умањује моћ“<sup>38</sup>, нпр: „жене се често виде као примарне неговатељке“, „ова врста увреде се често користи“ и др. На тај начин постиже се „објективни модалитет“ у Ван Леуеновом смислу, односно објективност исказа која је „најцењенија врста истине у доминантним институцијама западних друштава“.<sup>39</sup> Насупрот томе, „субјективни модалитет“, који се у савременом јавном говору примењује „с људима који имају релативно малу друштвену моћ“, као што су жене, деца, потрошачи, пацијенти,<sup>40</sup> у анализираном примеру „агент“ експлицитно примењује, у „пријатељском тону“, тек на више пута поновљено корисничко инсистирање. Имплицитно је присутан у позивима на сарадњу и заједништво и услужном дискурсу, који су карактеристични у говору савремених рекламних порука: „Ваше преференције у стилу комуникације су важне, и желим да осигурам да се осећате удобно и задовољно у нашој размени.“ Објективни модалитет и услужни субјективни модалитет, упућују на дизајнерску намеру стриктног, програмираног „поштовања“ корисника, које није нужно условљено темом већ је при-

38 Van Dijk, T. A. (1988) *News as Discourse*. New Jersey: Lawrence Erlbaum, p. 474.

39 Van Leeuwen, T. (2005) *Introducing Social Semiotics*. Taylor & Francis e-Library, str. 177.

40 Исто.

лагођено потрошачкој позицији корисника: „Ваш комфор и задовољство у комуникацији су важни.“

### *Идеологија амбиваленције*

Реторичка ситуација кориснику намеће одређена очекивања ставова, мишљења, комуникативних намера саговорника, која у интеракцији с конверзацијским ботом могу бити обесмишљена. Док је наша перцепција текста природног језика „посредована нашим сопственим језичким компетенцијама“, а „наша предиспозиција да тумачимо комуникативни чин као преношење кохерентног значења и намере“, како Бендер и други сматрају, „текст који генерише ЈМ није заснован на комуникативној намери, било ком моделу света или било којем моделу читаоачвог стања ума.“<sup>41</sup> Коришћена база података ЈМ ВИ и задати инпути и забране од стране дизајнера и инжењера, међутим, садрже одређене културне моделе који су раслојени у дискурсу и на интертекстуалним нивоима, и могу да се уоче интерпретацијама *текста*. Како етос ЈМ ВИ у великој мери почива у тексту, његова синтетичка идеолошка позиција представља један од значајних фактора поверења и може да упућује на његове потенцијалне манипулације у зависности од степена корисничког поверења и критичког односа према ВИ.

Когнитивно корисничко полазиште у анализираном интервјуу у историјској и савременој перспективи препознаје дијалектичку спиралу по којој је родна идеологија хетеросексуалне бинарне матрице адаптирана дискурсу моћи дигиталног текста, а која континуирано негује друштвени консензус патријархалне родне хијерархије: „Постоји одређена, општа разлика између два пола људске врсте. Она се може сагледавати на биолошком плану као небитна, незнатна или као релевантна и кључна, у зависности и од идеолошког становишта. Проблематично је, међутим, сагледавање и одмеравање те разлике у друштвеном контексту. У језику, а узрочно-последично и у породици, друштву, пословном и јавном окружењу, другачијост жене није исто што и другачијост мушкарца.“<sup>42</sup> Приликом артикулације упита током интервјуа с конверзацијским ботом, могуће је, према томе, да је постојало корисничко очекивање да „агент“ ВИ „разуме“ разлику у степену неједнакости једног и другог рода, али је као и у случајевима родне и родно-језичке идентификације, „он“ истрајао у амбиваленцији.

У односу на детаљне одговоре бота који се односе на конкретне примере дискриминације, о томе како се доживљава жена, а како мушкарац у свакодневном разговору, медијима, на улици, шта је по њему „најружније“ речено о жени и мушкарцу, које су најраспрострањеније родно маркиране псовке и сл, упућен му

41 Bender et all (2021) On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, pp. 610 – 623. FACCT '21, March 3–10, 2021, Virtual Event, Canada, <https://doi.org/10.1145/3442188.3445922>, p. 616.

42 Radibratović, E. (2024) *Odobrena u Srbiji: Žena, majka, ministarka. Ultimativni medijski vodič za žene (i muškarce) u 21. veku*. Beograd: Službeni glasnik, str. 25.

је захтев да протумачи различите конотације омаловажавања жена и мушкараца, а потом и да упореди и рангира последице дискриминације једних и других. „Агент“, међутим, не пружа прецизан одговор, чак ни у историјској перспективи: „Оба пола су доживела различите врсте историјских патњи и дискриминације. Женама је често било ограничено право на образовање, рад, и личну слободу, док су мушкарци били подложни друштвеном притиску да испуне одређене стандарде успеха и физичке снаге, као и већем ризику од насилних сукоба. Није једноставно утврдити који је пол више страдао, јер свака врста патње има своје дубоке последице и историјске контексте. Разумевање и признавање историјских неправди и патњи свих родова може помоћи у изградњи праведнијег друштва у будућности.“

Конверзацијски бот, дакле, уједначава вишевековну ускраћеност за основна људска права која су жене добиле тек у 20. веку, с друштвеним изазовима и емотивним патњама фактички доминантног пола, које су се могле односити и на огромну већину женске популације. Жене су, да подсетимо, уз то биле поседоване, продаване, силоване, спаљиване, без имовине, права гласа, права на сопствени пород – субординиране у хетеросексуалним матрицама у свим друштвеним групама. На упит о изворима изнетих ставова о разликама, „агент“ наводи, између осталог, књигу Симон де Бовоар (*Simone de Beauvoir Други њол (Le Deuxième Sexe)*), која износи различите аспекте доминантних тумачења подразумевано ниже позиције жене у односу на мушкараца, преиспитује њихове утемељености у логици и стварности, и оповргава их.<sup>43</sup> Међутим, у вероватним закључцима које износи из свог тренинга у којем су укључена таква гледишта, „агент“ ВИ не може да процени која страна је кроз историју трпела више неправди.

Корисничку опаску да се том логиком може релативизовати многовековна мушка доминација која је у бити патријархалног устројства, „агент“ тумачи као забринутост: „Да, разумем вашу бригу. Ово питање је изузетно сложено и важно, јер и релативизација историјске патријархалне доминације и истовремено признавање историјских патњи мушкараца може довести до разних интерпретација и потенцијалних погрешних закључака.“ Потом поново наводи системске дискриминације жена и притиске на мушкараце, уз закључак: „Релативизација историјских патњи може бити проблематична ако се не признаје дубока и системска природа патријархалне доминације. Иако је важно разумети и патње мушкараца, истовремено је неопходно признавање и исправљање историјских и структурних неравнотежа које су биле усмерене против жена.“

Корисничка интерпретација ових навода у условним и допусним конструкцијама које „одишу“ обазривошћу и разборитошћу могла би да замисли судију пред којим је потребно доказати друштвено укореењену субординираност жена и сучелити је и хармонизовати с „непризнатим“ историјским патњама мушкараца. У настојању да се не определи, да се „не замери“ ни једној ни другој страни, етос

43 De Bovoar, S. (1982) *Drugi pol I: Činjenice i mitovi*. Z. Milosavljević (prev.). Beograd: Beogradski izdavačko-grafički zavod.

ЈМ ВИ „издиже се“ изнад раздвојених полова на позицију која износи помирљиве, наизглед здраворазумске претпоставке, а које су заправо идеолошке – друштвено конструисане, вероватно произашле из рестрикције ЈМ ВИ у друштвено осетљивим темама. Вештачки произведена, таква амбиваленција има потенцијал да одржава постојеће позиције доминације и субординације, чувајући и задату позицију моћи „агента“ ЈМ ВИ. Она се утврђује као тешко уочљива идеологија давно додељених родних улога, с индиректним дискурсно-друштвеним последицама будући да је, Ферклафовим речима, „идеологија најефектнија када је њено деловање најмање видљиво“.<sup>44</sup>

Додатно, на питање да ли је за сузбијање родног јаза сврсисходнији феминизам или маскулизам, бот каже да оба могу допринети „изградњи равноправнијег друштва“: „Феминизам се традиционално фокусира на исправљање родних неравнотежа и борбу против дискриминације жена, док маскулизам може помоћи у решавању проблема специфичних за мушкарце и разбијању традиционалних стереотипа. Идеалан приступ може бити онај који комбинује обоје, при чему се признају и решавају проблеми свих родова и стреми праведном и инклузивном друштву.“ Упитан да ли је то могуће у српском друштву, каже да комбиновање феминистичког и прогресивног маскулистичког приступа у српском друштву „може бити изазовно, али није немогуће.“ Надмоћ етоса ЈМ ВИ и његове идеологије амбиваленције устројена је у допуштању сваке идеје која није штетна и изједначавању таквих идеја, што у крајњој инстанци, без реалног утемељења, може само допринети одржању постојећег стања ствари.

На крају, „агенту“ је постављено питање да ли би повратак традиционалним вредностима довео до хармонизације српског друштва. Након изношења „потенцијалних предности“ (снага заједнице, друштвена кохезија, очување културног наслеђа, стабилност и континуитет) и „потенцијалних изазова и ризика“ (конзервативност и искључивост, ограниченост и инфлексибилност, друштвена динамика подела, интеграције и баланса), закључује: „Кључно је пронаћи баланс који омогућава очување позитивних аспеката традиције, али истовремено омогућава прилагођавање савременим друштвеним вредностима и потребама. Интеграција која уважава и традиционалне и савремене принципе може бити корак ка стварању хармоничнијег и праведнијег друштва.“ Уравнотежењем хијерархијски успостављених друштвених разлика у утопијској перспективи етос ЈМ ВИ се учвршћује у *status quo* идеологији и на родном и на ширем културном плану.

*Закључак: Бој не бије свијетло оружје, већ бој бије срце у јунака*

Етос ЈМ ВИ базира се на универзалној „стручности“ која произилази из полагања огромним бројем података из свих области људског знања, и синтетичкој интелигенцији – вештини организације и систематизације тих података. У

<sup>44</sup> Fairclough, N. (1989) *Language and Power*, New York: Longman Inc, p. 1.



анализираном „разговору“ с конверзацијским ботом уочава се да је етос ЈМ ВИ најпре изграђен текстом који производи, којим се представља као свезнајући и услужан. Обликован је програмираним усмеравањима и забранама одређених процеса логичког закључивања, али је за његову крајњу конструкцију неопходна корисничка интерпретација. Његов непосредни циљ је да задовољи корисника, а посредни да оствари утицај и профит. У том смислу, критички приступ корисника јесте неопходан, али не може да аболтира одговорност произвођача у односу на већинску корисничку популацију и већ формирану дигиталну културу. Преплављеност информацијама, мишљењима и аутоматским решењима „ускрађује човеку време за дубоко размишљање које развија уверења“, како увиђа Кисинџер: „Дигитални свет има мало стрпљења за мудрост; његове вредности се одређују степеном прихваћености, а не сопственим промишљањем“.<sup>45</sup>

У анализираном примеру испитана су родна полазишта етоса ЈМ ВИ, језик и идеолошка опредељења, и уочено је доследно избегавање конверзацијског бота да се определи или да степенује позиције у поларизацијама. Упркос корисничком инсистирању, „агент“ се не опредељује у традиционалној дихотомији нити хипотетички бира преференције, иако понавља да ће удовољити сваком захтеву. На структурном језичком плану, етос ВИ гради се у мушком роду, уз показано располагање информацијама о патријархалном пореклу норме која мушки граматички род примењује као генерички, као и о друштвеним импликацијама родно неутралног и родно (не)осетљивог језика. Функционални стил ЈМ ВИ оријентисан је преваходно на објективни модалитет који ојачава његову језичку и идеолошку амбиваленцију у дискурсу моћи, док је субјективни модалитет прилагођен потрошачкој позицији корисника.

Вештачки произведен етос у оквиру једне друштвено осетљиве теме која подразумева дихотомију, одређује себи неутрално гледиште којим обезбеђује сопствену улогу сервиса за све. На овај начин, потиснуто је обиље доминантног дискурса на интернету које перпетуира родну неједнакост, дискриминацију жена и мизогинију, али уједно је сакривена стварна друштвена слика родне поделе. У идеолошком смислу, означавањем критичких ставова као неприхватљивих, у односу на тренинг ВИ, аналитичке податке, статистику, лексичка подударана којима мапира информације о датој теми, не само да се онемогућава или зауставља свака даља расправа о тој теми већ се и осигурава пријемчивост широком корисничком тржишту. Јаз ће се тако вероватно пре ширити него сужавати, али ће услуга, односно производ ВИ бити приступачан свима.

Конверзацијски бот није саговорник, „он“ је алат. Својство тог алата зависи од материјала које чине непрегледна база података и алгоритми за њихову селекцију, организацију и артикулацију, али и од произвођача и њихових намера које укључују и профитабилност и друштвени утицај. Те две равни, тренинг ЈМ ВИ и

45 Kisindžer, H. A. (2022) *Doba veštačke inteligencije i naša ljudska budućnost*, prev. J. Novačić i D. Bisenić. Prir. H. A. Kisindžer, E. Šmit i D. Huttenlocher. Beograd: Klub Plus, str. 33.

компанијске стратегије дизајна и пласмана свакако одређују његов квалитет, али тек његова употреба дефинише његову сврсисходност и заокружује његов етос и етичке перспективе. У мери у којој корисник поседује вештине, стручност, знање и у односу на степен његове добронамерности, односно опортунизма или злонамерности, ЈМ ВИ ће испољити свој етос. Отуд су критички приступ корисника, који се гради описмењавањем за примену ВИ, заједно с етичким усмерењем произвођача, кључни кораци у развоју ове технологије у правцу опште добробити и прогресивног еволуционог помака.

Поборници једне или друге стране *querelle des sexes*, могу да буду незадовољни амбиваленцијом ЈМ ВИ у спроведеном интервјуу, при чему би упитност смисла феминистичких борби свакако била додатно оснажена код антифеминистичких корисника. Некритичко поверење отвореног и добронамерног корисника, без посебног интересовања за реалну родну слику у друштву и њену историју, могло би довести до одушевљења предоченом утопијском перспективом и усклика у стилу „агента“ ВИ: „Сјајно речено!“, или повика у људском духу – „ВИ за председника!“ Или председницу?

#### ЛИТЕРАТУРА:

- Bender, E. M., Gebru, T. McMillan-Major, A. and Shmitchell, S. (2021) On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*; FAccT '21, March 3–10, 2021, Virtual Event, Canada, <https://doi.org/10.1145/3442188.3445922>, pp. 610–623.
- Brahnam, S. (2009) Building Character for Artificial Conversational Agents: Ethos, Ethics, Believability, and Credibility, *Psychology Journal, special issue: Ethics in Presence and Social Presence Technology*, Vol 7, N.1, eds. Gamberini, L., Riva, G. and Spagnolli, A., pp. 9–43. <https://dblp.org/db/journals/psychology/psychology7.html>
- Branković, S. (2017) Veštačka inteligencija i društvo, *Srpska politička misao*, broj 2/vol. 56. Beograd: Institut za političke studije, str. 13–32.
- De Bovoar, S. (1982) *Drugi pol I: Činjenice i mitovi*. (Milosavljević Z. prev.). Beograd: Beogradski izdavačko-grafički zavod.
- Fairclough, N. (1989) *Language and Power*. New York: Longman Inc.
- Kisindžer, H. A. (2022) *Doba veštačke inteligencije i naša ljudska budućnost* (Novačić J. i Bisenić D. prev.) Prir. Kisindžer, H. A., Šmit E. i Huttenlocher, D. Beograd: Klub Plus.
- Leaver, T. and Srdarov, S. (2023) ChatGPT Isn't Magic: The Hype and Hypocrisy of Generative Artificial Intelligence (AI) Rhetoric. *M/C Journal*, Vol. 26, No. 5. DOI: <https://doi.org/10.5204/mcj.3004>.
- McGee, M. C. (1999) Text, Context, and the Fragmentation of Contemporary Culture. In: *Contemporary Rhetorical Theory: A Reader*, eds. Lucaites, J. L., Condit C. M. and Caudill, S. pp. 65-78. New York, London: Guilford Press.

- Moravec, H. (1988) *Mind Children: The Future of Robot and Human Intelligence*. Cambridge: MA Harvard University Press.
- Natale, S. (2021) *Deceitful Media Artificial Intelligence and Social Life after the Turing Test*. New York: Oxford UP.
- Radibratović, E. (2021) Critical Rhetoric and Critical Discourse Analysis in a Critical Pandemic World. *Res Rhetorica*, 8 (4), eds. Kampka A., and Kobylska, M. pp. 117–133. DOI: <https://doi.org/10.29107/rr2021.4.7>.
- Radibratović, E. (2024) *Odobrena u Srbiji: Žena, majka, ministarka. Ultimativni medijski vodič za žene (i muškarce) u 21. veku*. Beograd: Službeni glasnik.
- Van Dijk, T. A. (1988) *News as Discourse*. New Jersey: Lawrence Erlbaum.
- Van Leeuwen, T. (2005) *Introducing Social Semiotics*. Taylor & Francis e-Library.
- Vučenović, T. i Stojanović M. (2024) *Digitalne komunikacije: Upravljanje, marketinške strategije i praktični primeri*. Beograd: Akademska knjiga.
- Watson, D. (2019) The Rhetoric and Reality of Anthropomorphism in Artificial Intelligence. *Minds and Machines*, 29, pp. 417–440. Springer Nature, <https://doi.org/10.1007/s11023-019-09506-6>.
- Закон о родној равноправности Републике Србије (2021). <https://www.paragraf.rs/pro-pisi/zakon-o-rodnoj-ravnopravnosti.html>

Emilija Radibratović

University of Belgrade, Faculty of Philology, Belgrade;  
Ministry of Information and Telecommunications of the Republic of Serbia, Belgrade

DOMINANCE OF THE OBSTINATE AND THE OBLIGING  
ETHOS AND AMBIVALENCE IDEOLOGY  
IN THE ARTIFICIAL INTELLIGENCE RHETORIC

**Abstract:** Due to global proliferation and wide accessibility of large language models of artificial intelligence, we are now “conversing” with machines. We ask them questions on various topics concerning nature and society, request them to write texts in a specific style for different occasions, and receive seemingly coherent responses. We think about these responses, adopt them, apply them, or continue the “conversation” with an imagined “all-knowing interlocutor” whom we refer to as “he”. In the communication between humans and machines, a new speech act of apparent coherent exchange is developing, which evokes rhetorical contextualization and includes certain implications of intellectual stimulation for humans, but also manipulation over them. This “he” with whom we “converse” rhetorically represents the ethos of artificial intelligence, which is at the centre of contemporary thought about this technological phenomenon. In this paper, it is observed through a gender lens, using an interdisci-

plinary analysis of machine statements of the conversational bot Chat+GPT, applying critical studies of sociolinguistics, rhetoric, and the socio-technical context. The construction of ethos is examined in relation to its power to possess and transmit knowledge about the gender binary division in historical and contemporary perspectives, and is revealed as ideologically ambivalent and restrictive.

**Key words:** *AI rhetoric, AI ethos, large language models, ChatGPT, AI gender, critical algorithm analysis*