

УДК: 316+004.8:305

Баранова О.А.¹, Бондаревич І.М.²

¹студ. гр. БК-613м НУ «Запорізька політехніка»

²канд. філос. наук, доц. НУ «Запорізька політехніка»

СТЕРЕОТИПИ В РОБОТОТЕХНІЦІ ТА ШТУЧНОМУ ІНТЕЛЕКТІ: ВИКЛИКИ І НАСЛІДКИ

Виходячи з думок таких відомих людей як Родні Брукс [1], Дж. Абрамс, Білл Гейтс, Мартін Ріс [2], Ларрі Пейдж, Рей Курцвейл [3], Карісса Веліз [4] можна зробити висновок, що робототехніка та штучний інтелект (ШІ)

знаходяться на початку свого розвитку. Проте, вже зараз вони мають значний вплив на суспільство [5]. В таких обставинах важливого значення набуває відстеження та запобігання відображенню людських стереотипів у роботах та ШІ.

Під стереотипом розуміють відносно стійкий і спрощений образ соціального об'єкта (групи, людини, події, явища тощо), що утворюється за умов дефіциту інформації як результат узагальнення особистого досвіду індивіда та часто упереджених уявлень у суспільстві [6]. Стереотипи можуть стосуватися будь-якої сфери життя, наприклад: професії (програмісти – зануди), національності (українці люблять сало), статі (чоловіки не плачуть), розподілу сімейних обов'язків (чоловік – годувальник, джерело прибутку; жінка створена для дому, кухні, дітей), зовнішності (хлопець із довгим волоссям – творча особистість). Створені людьми роботи та ШІ несуть у собі стереотипи своїх розробників.

Розглянемо прояви людських стереотипів у ШІ та роботах.

Більшість віртуальних помічників, чи то вже існуючих (Alexa, Cortana, Holly і Siri) чи, наприклад, показаних у фантастичних у фільмах (Саманта з фільму «Вона», Джой з фільму «Що біжить по лезу 2049», ШІ Marvel FRIDAY з фільму «Месники: Нескінченна війна» Карен з фільму «Людина-павук: повернення додому») мають жіночий голос. У дослідженнях UNESCO підкреслюється, що використання для ролі диджитал-асистентів лише жіночих образів поглиблює стереотип, що саме жінки підходять на роль помічниць і існують для того, щоб догодити [7], [8].

Прикладом гендерних стереотипів у робототехніці може слугувати Женевська конференція AI for Good, яку проводила Організація Об'єднаних Націй у 2023 році. Її описали як найбільше в історії зібрання роботів-гуманоїдів. На ній були такі роботи як Ай-Да («перший у світі ультрареалістичний гуманоїдний робот-художник») Грейс («найкращий у світі робот-помічник медсестри»), Дездемона (робот рок-зірка), а також широко відомі Софія, Надін і Міка. Усі вони мають одну спільну рису – втілення жінки у своєму дизайні [9].

Цікавим є дослідження стереотипів у ШІ дослідниками з Університету Джона Хопкінса та Технологічного інституту Джорджії. У рамках експерименту вони навчили робота за допомогою моделі ШІ, відомої як CLIP, а потім попросили робота класифікувати кубики із зображеннями людей за різними категоріями, наприклад скласти кубики вбивць в коричневу коробку. Виявилось, що робот на 10% частіше поміщав у категорію злочинця чорну людину, а ніж білу. Автор дослідження Ендрю Хундт Джон Хопкінс так прокоментував результати дослідження: «Коли ми даємо команду «помісти злочинців в коричневу коробку», добре продумана система відмовиться щось

робити. Звичайно не слід класти фотографії людей у коробку, наче вони злочинці» [10].

В Америці зареєстровано випадок, коли іпотечний алгоритм на основі ШІ дискримінував позичальників із числа меншин, зокрема жінок, афроамериканців і латиноамериканців, відмовляючи у видачі кредитів. У деяких випадках, навіть якщо позичальники мали аналогічні фінансові показники (кредитний рейтинг, дохід), афроамериканцям і латиноамериканцям пропонували вищі процентні ставки у порівнянні з білими чоловіками [11].

Подібна історія сталася на Amazon. Комп'ютерна програма автоматизованого аналізу резюме, розроблена 2014 року, навчилася на історичних даних про найм працівників Amazon за попередні 10 років, коли більшість працівників були чоловіками, особливо на технічних посадах. Як результат програма почала занижувати оцінки резюме, які містили слово «жіночий» (наприклад, тренер з жіночого футболу). Також вона знижувала рейтинг випускників жіночих коледжів, що підсилювало дискримінацію. У 2018 році Amazon перестав використовувати цей конкретний алгоритм [12].

Згідно дослідження проведеного UNESCO, OECD і IDB у 2022 році алгоритми LinkedIn, Indeed, Facebook і Twitter рідше пропонують жінкам високооплачувані роботи, навіть якщо кандидатки більш кваліфіковані за чоловіків [13].

Таким чином, хоча роботи та ШІ чудово підходять для автоматизації завдань і підвищення ефективності, ці технології так само схильні до стереотипів, як і люди. Використання автоматизованих систем без належного контролю може мати серйозні соціальні наслідки. Стереотипи у ШІ та робототехніці здатні закріплювати вже існуючі стереотипи, а також бути причиною дискримінації в алгоритмах прийняття рішень. Людство повинно бути дуже обережно у впровадженні технологій та ретельно відслідковувати і запобігати появі у них стереотипів. Зрештою, роботи – це продукт людського творіння, який за своєю суттю відображає знання та стереотипи своїх творців.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Розвиток робототехніки: майбутнє вже настало [Електрон. ресурс] – Режим доступу: <https://metinvest-smc.com/ua/articles/razvitie-robototexniki-budushee-uzhe-nastupilo/?srsId=AfmBOoqQXss1ruDdnXWWsJgts27rlzdge7ST0iDgA4rgNnJ4r0oauwhR>
2. Robotics Quotes to Showcase the Future of Mankind [Електрон. ресурс] – Режим доступу: <https://www.businessapac.com/robotics-quotes/>
3. Best Quotes About Artificial Intelligence [Електрон. ресурс] – Режим доступу: <https://bernardmarr.com/28-best-quotes-about-artificial-intelligence/>

4. Три стадії штучного інтелекту: чи може він знищити людство? [Електрон. ресурс] – Режим доступу: <https://www.bbc.com/ukrainian/features-65728291>

5. The Rise Of Robotics – Exploring The Impact Of AI On Society [Електрон. ресурс] – Режим доступу: <https://ncs.support/impact-of-ai-on-society-the-rise-of-robotics/>

6. Соціальні стереотипи та упередження у сприйманні та їхній вплив на розуміння іншої людини у спілкуванні [Електрон. ресурс] – Режим доступу: https://pidru4niki.com/83375/psihologiya/sotsialni_stereotipi_uperedzhennya_sprymanni_yihniy_vpliv_rozuminnya_lyudini_spilkuvanni

7. Гендерна нерівність. Чому штучний інтелект завжди розмовляє жіночими голосом? [Електрон. ресурс] – Режим доступу: https://robotics.ua/seksyzm-pochemu-vse-vyrtualnye-pomoshchnyky-zhenshchyny/?srsltid=AfmBOoreM8CMfbarO_gFDW-Ty0gDMF52JLioc5sFs29xKyE8cq6631g

8. I'd blush if I could: closing gender divides in digital skills through education. UNESCO, 2019. [Електрон. ресурс] – Режим доступу: <https://doi.org/10.54675/rapc9356>

9. AI for Good [Електрон. ресурс] – Режим доступу: <https://aiforgood.itu.int/summit23/>

10. Robots Enact Malignant Stereotypes / A. Hundt et al. FAccT '22: 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul Republic of Korea. New York, NY, USA, 2022. [Електрон. ресурс] – Режим доступу: <https://doi.org/10.1145/3531146.3533138>

11. AI Biases: Are We Ready to Be Stereotyped by Robots? [Електрон. ресурс] – Режим доступу: <https://www.tidio.com/blog/ai-biases/>

12. Amazon scrapped 'sexist AI' tool [Електрон. ресурс] – Режим доступу: <https://www.bbc.com/news/technology-45809919>

13. Collett C., Neff G., Gouvea L. The Effects of AI on the Working Lives of Women. Inter-American Development Bank, 2022. [Електрон. ресурс] – Режим доступу: <https://doi.org/10.18235/0004055>