

5) У структурі романтичних взаємин довіра виконує функцію психологічного «буфера», який пом'якшує інтенсивність конфлікту, підтримує суб'єктивне відчуття безпеки та створює умови для діалогу. Саме через довіру конфлікт може трансформуватися з деструктивного зіткнення позицій у ресурс розвитку стосунків і поглиблення взаєморозуміння (Чуйко, Чаплак, 2016).

ЛІТЕРАТУРА

1. **Березовська Л. І., Юрков О. С.** Психологія конфлікту : навчально-методичний посібник для студентів напряму підготовки «Практична психологія». Мукачєво : Мукачівський державний університет, 2016. 201 с.
2. **Бігунов О. В.** Психологічні особливості міжособистісних стосунків у ранній дорослості. Психологічні перспективи. 2018. № 31. С. 23–32.
3. **Гірник А. М.** Основи конфліктології : навчальний посібник. Київ : Академвидав, 2010. 320 с.
4. **Лисенко Л. А.** Довіра як психологічний феномен. Науковий вісник Херсонського державного університету. Серія «Психологічні науки». 2016. Вип. 3. С. 41–46.
5. **Чуйко О. В., Чаплак Я. М.** Міжособистісна довіра як психологічний феномен. Проблеми сучасної психології. 2016. Вип. 34. С. 312–321.

ЧАЙКОВСЬКА ДАР'Я МИХАЙЛІВНА, студентка МП-1«Психологія», НаУКМА,
ГОЛУБЄВА МАРІЯ ОЛЕКСАНДРІВНА, доцент, кандидат педагогічних наук,
доцент кафедри психології та педагогіки, НаУКМА

ГЕНДЕРНА УПЕРЕДЖЕНІСТЬ АЛГОРИТМІВ ШТУЧНОГО ІНТЕЛЕКТУ ЯК ПСИХОЛОГІЧНИЙ ЧИННИК СОЦІАЛЬНОЇ НАПРУГИ ТА КОНФЛІКТІВ

Вступ. Штучний інтелект дедалі активніше використовується у сферах ухвалення соціально значущих рішень, зокрема під час найму та відбору персоналу, у фінансових, медичних, судових та медійних практиках. У цих контекстах алгоритмічні системи виступають не лише технічним інструментом, а посередником між індивідом і соціальними інституціями, впливаючи на доступ до ресурсів, можливостей і соціального статусу (переклад з англійської тут та далі авторки тез – Д. Ч.) (Nadeem, Abedin, Marjanovic, 2020).

Попри уявну нейтральність, алгоритмічні рішення здатні відтворювати й підсилювати соціальні нерівності, зокрема гендерні, маскуючи їх під виглядом об'єктивних закономірностей. Із психологічної точки зору взаємодія з упередженими системами впливає на соціальне сприйняття, міжгрупові установки та переживання справедливості, формуючи відчуття безсилля перед автоматизованими процесами (Ho, Hartanto, Koh, Majeed, 2025). Накопичення таких ефектів, феноменів та процесів сприяє зростанню соціальної напруги, зниженню довіри до інституцій і загостренню міжгрупових суперечностей. Саме тому, тема нашої наукової розвідки «Гендерна упередженість алгоритмів

штучного інтелекту як психологічний чинник соціальної напруги та конфліктів» є досить актуальною.

Виклад основного матеріалу. На початку нашого аналізу звернемося до соціально-психологічних засад гендерної упередженості алгоритмів штучного інтелекту (далі - ШІ). У соціальній психології гендерні упередження розглядаються як наслідок соціалізації та засвоєння стереотипів, які формують стійкі уявлення про соціальні групи й впливають на сприйняття інших людей. В умовах цифровізації алгоритмічні системи стають новим каналом поширення таких уявлень, відтворюючи культурно закріплені норми у масштабованій і стандартизованій формі (Nadeem, Abedin, Marjanovic, 2020).

Ми приєднуємося до думки про те, що гендерна упередженість у штучному інтелекті визначається як систематичне викривлення алгоритмічних результатів, що відтворює або підсилює наявні в суспільстві гендерні стереотипи та нерівності. Такий різновид упередженості проявляється у нерівній представленості соціальних груп, асиметриях алгоритмічного ранжування та оцінювання, а також у стабільних відмінностях результатів для різних гендерних груп. Це створює відповідні сталі викривлення. Автоматизований і повторюваний характер таких викривлень робить їх менш помітними, але більш стійкими, маскуючи соціальне походження нерівності під виглядом технічної закономірності (Nadeem, Abedin, Marjanovic, 2020).

У межах огляду досліджень виділяють низку ключових чинників формування гендерної упередженості в алгоритмах штучного інтелекту, які охоплюють як соціальний контекст використання технологій, так і особливості їхньої технічної реалізації (Nadeem, Abedin, Marjanovic, 2020; Ho, Hartanto, Koh, Majeed, 2025).

Узагальнюючи ці підходи, Д. Чайковська прийшла до висновку про те, що гендерна упередженість формується на трьох взаємопов'язаних рівнях:

- Навчальні дані. Алгоритмічні системи навчаються на історичних масивах інформації, що відображають практики дискримінації та нерівного доступу до ресурсів, закладаючи соціальні асиметрії в майбутні алгоритмічні рішення.

- Архітектура та алгоритмічний дизайн. Проектні рішення, вибір змінних і критерії оптимізації відображають ціннісні припущення розробників і інституційний контекст створення технологій, унеможливаючи повну нейтральність алгоритмів.

- Петлі зворотного зв'язку з користувачами. Алгоритмічні результати впливають на поведінку користувачів, а ці реакції потрапляють у нові дані, сприяючи закріпленню домінантних соціальних патернів і самопідтримуваному характеру гендерної упередженості.

На цих рівнях реалізуються такі чинники, як брак різноманіття в даних і серед розробників, суспільна гендерна упередженість, мовна дискримінація, а також економічні й поведінкові стимули, що опосередковано підсилюють нерівність (Nadeem, Abedin, Marjanovic, 2020).

Психологічно значущим є механізм самопідтримуваного відтворення гендерної нерівності, опосередкований алгоритмічними системами. Дослідження низки сучасних науковців свідчать, що наявна соціальна гендерна нерівність відображається в алгоритмічних репрезентаціях, зокрема в результатах інтернет-пошуку, а регулярна експозиція до таких результатів сприяє формуванню в користувачів упереджених когнітивних прототипів щодо «типових» соціальних ролей. Сформовані когнітивні прототипи, у свою чергу, впливають на очікування та поведінкові рішення людей, зокрема у сфері найму й професійного відбору, що призводить до відтворення початкової нерівності в соціальних практиках. У такий спосіб вибудовується замкнений психологічний механізм, у межах якого соціальна гендерна нерівність через алгоритмічні репрезентації трансформується в індивідуальні когнітивні уявлення та закріплюється через рутинні поведінкові рішення (Vlasceanu, Amodio, 2022).

Дослідження генеративних візуальних і мовних систем штучного інтелекту показують, що алгоритмічна упередженість проявляється як у нерівномірній представленості соціальних груп, так і у відтворенні стереотипних характеристик, знижуючи чутливість до дискримінації за умови сприйняття ШІ як об'єктивного джерела (Messingschlager, Appel, 2025). Окремі дослідження показують, що люди переносять звичні гендерні стереотипи й на самі системи штучного інтелекту. Залежно від того, як «позначений» алгоритмічний агент, змінюється рівень довіри до нього, готовність співпрацювати та очікування щодо його компетентності, що безпосередньо впливає на поведінку користувачів (Bazazi, Karpus, Yasseri, 2025).

Наведені вище об'єктивні процеси прямо ведуть від алгоритмічної упередженості до соціальної напруги та конфліктів.

Алгоритмічна гендерна упередженість набуває психологічного значення не лише через нерівну репрезентацію соціальних груп, але й через те, як результати роботи систем ШІ сприймаються, інтерпретуються та інтеріоризуються користувачами. Взаємодія з алгоритмічними рішеннями відбувається в ситуації, при якій користувачі/користувачки майже не мають можливості впливати на роботу систем або оскаржувати їхні рішення. Така нерівність у контролі й відповідальності формує психологічні реакції на сприйняту несправедливість і відчуття обмеженого впливу.

Одним із важливих психологічних чинників взаємодії з системами штучного інтелекту є тривожність щодо цих технологій. Дослідження показують, що жінки в середньому відчують більше занепокоєння стосовно ШІ, мають менш позитивне ставлення до нього, рідше використовують такі системи та оцінюють свої знання про них як нижчі порівняно з чоловіками (Russo, Romano, Clemente, Iacovone, Gladwin, Panno, 2025). Водночас гендерні відмінності у ставленні до ШІ зберігаються навіть тоді, коли рівень тривожності є низьким, що свідчить про вплив ширших соціальних і структурних чинників, а не лише емоційних реакцій. За високого рівня тривожності ці відмінності частково зменшуються, однак не зникають повністю, що вказує на обмежену роль тривожності у подоланні гендерного розриву.

Психологічні наслідки алгоритмічної упередженості посилюються в організаційному контексті. Дослідження сприйняття алгоритмічної упередженості показують, що системи штучного інтелекту, які використовуються організаціями, сприймаються людьми як більш загрозливі, ніж алгоритмічні інструменти індивідуального користування. Водночас користувачі відчувають значно менші можливості впливу на такі системи або зміни їхніх рішень (Overbye-Thompson, Rice, 2025). Поєднання високого відчуття ризику з переконанням, що люди майже не можуть впливати на роботу таких систем, не стимулює активні дії, а радше призводить до уникання, пасивного прийняття та психологічного дистанціювання. У такій ситуації алгоритмічна упередженість переживається як психологічна, соціальна й політична загроза, а не як окрема технічна помилка.

Особливо сильний психологічний вплив мають візуальні та медійні репрезентації, створювані системами ШІ. Емпіричні дані в онлайн-медіа та великих мовних моделях свідчать про стійке викривлення в гендерному та віковому аспектах: жінки систематично представлені як молодші, а чоловіки – як старші, незалежно від реальних демографічних показників (Guilbeault, Delecourt, & Desikan, 2025, с. 1129 – 1135). Такі викривлення посилюються в контекстах високого статусу, престижу та доходу, зокрема у професіях, пов'язаних із владою, експертизою та технологіями. Алгоритми ШІ не лише відображають наявні соціальні патерни, а й підсилюють їх, формуючи уявлення про «нормальний» вік, бажаність і легітимність кандидатів/кандидаток для престижних ролей.

Повторюваний контакт зі стереотипізованими репрезентаціями впливає на очікування та рішення людей, зокрема у сфері найму й професійного відбору, що підтверджує наявність реальних психологічних та інституційних наслідків алгоритмічної упередженості. На нашу думку, у сукупності такі процеси створюють феномен гендерної упередженості ШІ як психологічний механізм переходу від структурної нерівності до соціальної напруги. Алгоритмічні репрезентації формують упереджені когнітивні уявлення, які поєднуються з підвищеною тривожністю, відчуттям безсилля та переконанням, що люди не мають реальних можливостей впливати на рішення систем, створюючи умови для латентного конфлікту, недовіри до інституцій і загострення міжгрупових суперечностей.

Висновки. Гендерна упередженість у штучному інтелекті – сучасна психологічно-інформаційна категорія, яка визначає системне та систематичне викривлення алгоритмічних результатів, що відтворює або підсилює наявні в суспільстві нерівноправні гендерні стереотипи.

Гендерна упередженість алгоритмів штучного інтелекту є складним соціально-психологічним явищем, а не лише технологічною чи етичною проблемою. Сучасні алгоритмічні системи ШІ, інтегровані у процеси ухвалення соціально значущих рішень, відтворюють і легітимізують гендерну нерівність, надаючи їй вигляду об'єктивної закономірності. Автоматизований характер таких рішень сприяє їхній нормалізації та знижує критичну чутливість до

дискримінаційних наслідків.

Психологічні ефекти алгоритмічної упередженості ШІ проявляються через вплив на соціальне сприйняття, когнітивні уявлення та поведінкові рішення. Взаємодія з системами ШІ поєднується з тривожністю, відчуттям втрати контролю та низькою сприйнятою ефективністю впливу, особливо в організаційному контексті. У сукупності це формує механізм відтворення нерівності, що підсилює латентну соціальну напругу та недовіру до інституцій у особистості, що розмірковує критично.

Подальші психологічні дослідження мають бути спрямовані на вивчення індивідуальних і групових відмінностей у сприйнятті алгоритмічних рішень, ролі емоційних чинників і механізмів формування довіри до ШІ. Важливим є також пошук психологічних і освітніх інтервенцій, здатних послаблювати упереджені ефекти алгоритмів і сприяти більш рефлексивній та справедливій взаємодії людини з технологіями.

Література

1. Bazazi S., Karpus J., Yasseri T. AI's assigned gender affects human-AI cooperation. *iScience*. 2025. Vol. 28. Art. 113905. DOI: <https://doi.org/10.1016/j.isci.2025.113905>

2. Guilbeault D., Delecourt S., Desikan B. S. Age and gender distortion in online media and large language models. *Nature*. 2025. Vol. 646. P. 1129–1135. DOI: <https://doi.org/10.1038/s41586-025-09581-z>

3. Ho J. Q. H., Hartanto A., Koh A., Majeed N. M. Gender biases within artificial intelligence and ChatGPT: Evidence, sources of biases and solutions. *Computers in Human Behavior: Artificial Humans*. 2025. Vol. 4. Art. 100145. DOI: <https://doi.org/10.1016/j.chbah.2025.100145>

4. Messingschlager T. V., Appel M. Algorithmic bias in image-generating artificial intelligence: Prevalence and user perceptions. *Information, Communication & Society*. 2025. Advance online publication. DOI: <https://doi.org/10.1080/1369118X.2025.2584146>

5. Nadeem A., Abedin B., Marjanovic O. Gender bias in AI: A review of contributing factors and mitigating strategies. *Proceedings of the Australasian Conference on Information Systems (ACIS 2020)*. Wellington, New Zealand, 2020. URL: <https://aisel.aisnet.org/acis2020/27>

6. Overbye-Thompson H., Rice R. E., та ін. Perceptions of algorithmic bias: Risk and efficacy across individual and organizational uses. *International Journal of Human-Computer Interaction*. 2025. DOI: <https://doi.org/10.1080/10447318.2025.xxxxxx>

7. Russo C., Romano L., Clemente D., Iacovone L., Gladwin T. E., Panno A. Gender differences in artificial intelligence: The role of artificial intelligence anxiety. *Frontiers in Psychology*. 2025. DOI: <https://doi.org/10.3389/fpsyg.2025.xxxxx>

8. Vlasceanu M., Amodio D. M. Propagation of societal gender inequality by internet search algorithms. *Proceedings of the National Academy of Sciences of the United States of America*. 2022. Vol. 119, no. 29. Art. e2204529119. DOI: <https://doi.org/10.1073/pnas.2204529119>