

ВІДТВОРЕННЯ ГЕНДЕРНИХ СТЕРЕОТИПІВ У СИСТЕМАХ МАШИННОГО ПЕРЕКЛАДУ (НА МАТЕРІАЛІ УКРАЇНСЬКОЇ ТА АНГЛІЙСЬКОЇ МОВ)

REPRODUCTION OF GENDER STEREOTYPES IN MACHINE TRANSLATION SYSTEMS (BASED ON THE MATERIAL OF THE UKRAINIAN AND ENGLISH LANGUAGES)

Мороз М.Ю.,

orcid.org/0009-0007-0083-6501

аспірантка кафедри германської філології

Київського національного лінгвістичного університету

У статті проаналізовано вияви гендерної упередженості в машинному перекладі, спричинені специфікою алгоритмів штучного інтелекту та особливостями навчальних корпусів. Особливу увагу приділено тому, як автоматизовані системи перекладу відтворюють і закріплюють гендерні стереотипи, особливо під час перекладу між мовами, що мають відмінні граматичні системи. Розглянуто специфіку передачі гендерно-маркованої лексики, зокрема назв професій і посад, які залежно від контексту та статі особи можуть перекладатися за допомогою фемінітивів або маскулінитивів, враховуючи необхідність підкреслення гендерної специфіки.

У рамках аналізу гендерної упередженості в системах машинного перекладу було встановлено, що її формування обумовлено комплексом лінгвотехнологічних та соціокультурних факторів. Першочергове значення має антропоцентрична природа навчальних даних, які, будучи відображенням історичного мовного вжитку, інкорпують традиційні гендерні стереотипи та асиметричні моделі соціальних взаємин. Сукупність таких даних формує систематичне відхилення в алгоритмічній обробці мови. Визначено ключові виклики, що виникають під час перекладу нейтральних гендерних форм, а також проблеми забезпечення точності й стилістичної адекватності перекладеного тексту.

У статті розглянуто сучасні підходи до подолання гендерної упередженості в машинному перекладі, зокрема методи вдосконалення моделей для врахування гендерного контексту, впровадження додаткових анотацій у навчальні корпуси та застосування алгоритмів, що дозволяють користувачам обирати бажаний гендерний варіант перекладу. Особливий акцент зроблено на аналізі впливу культурних і мовних особливостей на процес автоматичного перекладу, а також запропоновано стратегії, спрямовані на мінімізацію гендерних стереотипів.

У статті також розглядається соціокультурний контекст, у межах якого гендерна упередженість виявляється через явище гендерної сегрегації на ринку праці. Проаналізовано ключові причини цього явища, серед яких виокремлено поширені гендерні стереотипи, економічні чинники, соціальні ролі, історично сформовані обмеження та особливості мовних практик. Показано, як мовні норми та традиції віддзеркалюють і водночас закріплюють гендерну сегрегацію на ринку праці. Наголошується, що процес адаптації мови до принципів гендерної рівності є невід'ємною складовою ширших соціокультурних трансформацій, спрямованих на подолання гендерних нерівностей у суспільстві.

Дослідження підкреслює важливість інтеграції принципів гендерної рівності в процеси розробки та вдосконалення систем машинного перекладу. Виокремлено ключові напрями для подальших наукових розвідок, серед яких: розширення навчальних корпусів текстами, що враховують гендерну інклюзивність; впровадження адаптивних алгоритмів для корекції гендерної упередженості; а також аналіз впливу технологій машинного перекладу на формування суспільних уявлень про гендерні ролі.

Ключові слова: гендерна упередженість, машинний переклад, гендерні стереотипи, гендерно-чутлива лексика, гендерно-маркована лексика.

The article examines the manifestations of gender bias in machine translation, arising from the specific characteristics of artificial intelligence algorithms and the composition of training corpora. Particular attention is paid to how automated translation systems reproduce and reinforce gender stereotypes, especially when translating between languages with distinct grammatical structures. The study examines the specifics of rendering gender-marked vocabulary, particularly occupational titles and positions, which may be translated using either feminine or masculine forms depending on the context and the gender of the individual, while considering the necessity of emphasizing gender specificity.

In the analysis of gender bias in machine translation systems, it has been established that its formation is driven by a combination of linguistic, technological, and sociocultural factors. A primary factor is the anthropocentric nature of training data, which, as a reflection of historical language usage, incorporates traditional gender stereotypes and asymmetric models of social interaction. The accumulation of such data leads to systematic biases in algorithmic language processing. Key challenges associated with translating gender-neutral forms and ensuring the accuracy and stylistic adequacy of the translated text are identified.

The article explores contemporary approaches to mitigating gender bias in machine translation, including methods for improving models to account for gender context, incorporating additional annotations into training corpora, and implementing algorithms that allow users to select their preferred gender variant of the translation. Special emphasis is placed on analyzing the impact of cultural and linguistic factors on the automatic translation process, as well as proposing strategies aimed at minimizing gender stereotypes.

The study also examines the sociocultural context in which gender bias manifests through the phenomenon of gender segregation in the labor market. Key causes of this phenomenon are analyzed, including widespread gender stereotypes, economic factors, social roles, historically established constraints, and linguistic practices. The article demonstrates how linguistic norms and traditions both reflect and perpetuate gender segregation in the labor market. It is emphasized that adapting language to the principles of gender equality is an integral part of broader sociocultural transformations aimed at overcoming gender inequalities in society.

The research underscores the importance of integrating gender equality principles into the development and improvement of machine translation systems. Key directions for further scholarly inquiry are outlined, including expanding training corpora with texts that account for gender inclusivity, implementing adaptive algorithms to correct gender bias, and analyzing the impact of machine translation technologies on shaping societal perceptions of gender roles.

Key words: gender bias, machine translation, gender stereotypes, gender-sensitive language, gender-marked language.

Постановка проблеми. Питання гендерної упередженості у машинному перекладі є актуальним викликом у сфері обробки природної мови та штучного інтелекту, що виникає через низку факторів, які потребують ретельного дослідження та вирішення.

Сучасні системи машинного перекладу базуються на статистичних моделях і великих обсягах текстових даних, які відображають існуючі гендерні стереотипи та упередження, присутні в суспільстві [1]. Як результат, моделі перекладу – алгоритмічні підходи до автоматизованого перетворення тексту з однієї мови на іншу, не лише засвоюють ці упередження, але й відтворюють їх у своїх перекладах, що впливає на якість і гендерну нейтральність отриманого тексту [2]. Крім того, граматичні системи різних мов демонструють значні відмінності у способах кодування гендерної інформації, що створює специфічні виклики для перекладачів [3]. Ця проблема особливо гостро постає при перекладі з мов з менш вираженою гендерною диференціацією, як-от англійська, на мови з більш чіткою гендерною системою, такі як українська. У процесі перекладу виникають труднощі з визначенням роду іменників та займенників, особливо коли контекст не надає достатньо інформації про стать особи [4, с. 57–61].

Так, зазначається, що системи машинного перекладу асоціюють професії та ролі з певним гендером, відтворюючи усталені стереотипи, що може призводити до спотворення змісту перекладу [5, с. 1679–1680]. Наприклад, “*nurse*” (медсестра/медбрат) сприймається та перекладається як жіноча професія, а “*doctor*” (лікар/лікарка) – як чоловіча, незалежно від фактичної статі особи, що призводить до посилення соціальної нерівності, неточності перекладів і підіймає етичні питання щодо політичної коректності та гендерної рівності в технологіях штучного інтелекту.

Аналіз останніх досліджень і публікацій. Сучасні дослідження в галузі машинного перекладу (МП) та обробки природної мови (ОПМ)

приділяють значну увагу проблематиці гендерної упередженості, зокрема, вивчають її вияви у нейронному машинному перекладі (НМП) та причини її виникнення [6]; взаємозв'язок між мовними структурами, соціокультурними стереотипами та системними упередженнями у машинному перекладі [7]; вплив гендерної упередженості на роботу сучасних моделей штучного інтелекту, таких як GPT-3 і ChatGPT [8]; методи оцінки гендерної упередженості в машинному перекладі (МП) на матеріалі різних мовних пар [5]; стратегії зменшення гендерної упередженості в МП, а саме створення гендерно-збалансованих навчальних даних для моделей НМП [9]; кореляцію між контекстною інформацією та рівнем гендерної упередженості з використанням інноваційних методів обробки лінгвістичних даних [10].

Попри значну кількість проведених досліджень, проблема гендерної упередженості в машинному перекладі залишається недостатньо вивченою, особливо в контексті стрімкого розвитку технологій штучного інтелекту. З огляду на більшу увагу до питань гендерної рівності та необхідність впровадження етичних принципів у мовні технології, ця проблема набуває особливої актуальності. Гендерна упередженість у системах машинного перекладу не лише негативно впливає на якість перекладу, але й сприяє відтворенню соціальних стереотипів, що посилює нерівність як у професійній діяльності, так і в суспільному житті загалом.

Метою нашого дослідження є системний аналіз випадків гендерної упередженості в системах машинного перекладу на матеріалі англійської та української мов, зокрема оцінка точності перекладу гендерно-маркованих одиниць, виявлення впливу таких упереджень на якість перекладу, а також визначення їхньої ролі у відтворенні та посиленні гендерних стереотипів.

Виклад основного матеріалу. Гендерна упередженість – це форма дискримінації, яка ґрунтується на стереотипних уявленнях про ролі й характеристики, притаманні людям певної статі та виявляється у необ'єктивному ставленні або пове-

дінці щодо особи чи групи осіб на основі їхньої статі, а не індивідуальних якостей чи здібностей [11]. У контексті машинного перекладу гендерна упередженість – це систематична помилка, яка виникає, коли системи МП відтворюють або навіть посилюють гендерні стереотипи та упередження, присутні в навчальних даних [12]. Вона часто виникає при перекладі з мов, які не мають граматичного роду (наприклад, англійської), на мови, де рід є обов'язковим граматичним маркером (наприклад, українська, німецька, французька) [1]. У таких випадках система машинного перекладу змушена обирати між чоловічим та жіночим родом, не маючи достатньо контекстної інформації. Дослідження показують, що системи машинного перекладу мають тенденцію використовувати чоловічий рід як стандартний варіант при перекладі гендерно-нейтральних слів, що є виявом явища відомого як «чоловічий рід за замовчуванням» (англ. male default) [13].

Одними з основних виявів гендерної упередженості в системах машинного перекладу є стереотипне приписування гендеру професіям та стереотипне приписування гендеру особам на основі додаткових характеристик, коли гендерні упередження виявляються у використанні прикметників для опису чоловіків і жінок, що відображає закріплені в суспільстві стереотипи. У рамках даного дослідження ми зосередимо увагу на аналізі стереотипного присвоєння гендеру професіям як одного з ключових аспектів гендерної упередженості в системах машинного перекладу.

Цей феномен виявляється у тенденції автоматичного асоціювання професій із певним гендером, що відображає традиційні соціальні ролі та стереотипи, закріплені в суспільстві. Наприклад, назви професій, таких як *“doctor”, “engineer”, “professor”, “police officer”, “firefighter”, “pilot”, “architect”, “lawyer”, “soldier”, “judge”, “driver”, “mechanic”, “builder”, “programmer”, “physicist”* зазвичай перекладаються у чоловічому роді, що підкреслює сприйняття цих професій як «традиційно чоловічих». Професії *“nurse”, “babysitter”, “cleaner”, “secretary”, “teacher”, “librarian”, “housekeeper”, “florist”, “beautician”* отримують жіночий рід у перекладі, відображаючи стереотипну асоціацію з «традиційно жіночими» ролями.

Такий мовний феномен є відображенням ширшого явища гендерної сегрегації на ринку праці. Так, розподіл професій на так звані «чоловічі» (Male-Dominated Occupations) та «жіночі» (Female-Dominated Occupations) є не лише лінгвістичним, а й соціально-економічним виявом, що виникає внаслідок низки ключових факторів:

1. Гендерні стереотипи

Однією з головних причин гендерного розподілу професій є стереотипи щодо ролей чоловіків і жінок у суспільстві. Жінкам традиційно приписують «доглядові» функції, що впливає на їхній вибір професій у сферах освіти чи медицини. Чоловіки ж частіше обирають технічні чи фізично важкі спеціальності через очікування суспільства щодо їхньої сили, витривалості та раціональності [14].

2. Економічні фактори

Гендерна нерівність в економічній сфері виявляється через сегрегацію на ринку праці, розрив в оплаті праці та нерівний доступ до ресурсів, що охоплює горизонтальний поділ (переважна кількість жінок у низькооплачуваних сферах діяльності) та вертикальний (домінування чоловіків на керівних посадах) і впливає на середню заробітну плату: жінки в Україні заробляють на 18–22% менше за чоловіків через недостатню представленість у високоприбуткових галузях і «штраф за материнство» – втрату кар'єрних можливостей через догляд за дітьми [15].

3. Соціальні ролі

Соціальні ролі визначаються як набір очікувань, норм і обов'язків, які суспільство покладає на індивідів залежно від їхньої статі чи соціального статусу [16, с. 27–29]. Вони формуються під впливом культурних традицій, сімейного виховання, суспільних стереотипів та економічних умов. У патріархальних суспільствах, зокрема в Україні, історично склалося, що чоловіки асоціюються з ролями годувальників і захисників, а жінки виконують ролі матерів та домогосподарок, що обмежує їхні кар'єрні можливості та впливає на вибір професій [17].

4. Історичні обмеження

Історичний контекст також сприяв формуванню гендерного розподілу у сфері зайнятості. Зокрема, в Україні з 1993 до 2017 року діяла заборона жінкам працювати на ливарних роботах, у сфері металургії, електротехнічного та хімічного виробництва через шкідливі чи небезпечні умови праці. Це обмежувало доступ жінок до певних галузей і закріплювало стереотипи про «чоловічу» природу цих сфер [18].

5. Мовні фактори

Проблема гендерної сегрегації має й лінгвістичний вимір, який виявляється у використанні гендерно-упередженої термінології та мовних практик, що сприяють закріпленню нерівності між чоловіками і жінками на соціальному, культурному та професійному рівнях.

Так, в англійській мові чітка гендерна диференціація в назвах професій, ролей і соціальних функцій, традиційно виявлялася через викори-

стання специфічних суфіксів і слів, що відображало патріархальні уявлення про гендерні ролі, закріплені в суспільстві. Зокрема, для позначення жіночої статі використовувалися суфікси *-ess*, *-euse* та *-trix*, серед яких суфікс *-ess* був одним із найпоширеніших та слугував засобом утворення жіночих варіантів слів, наприклад: *authoress* (письменниця) на противагу *author* (письменник), *poetess* (поетеса) замість *poet* (поет) або *actress* (актриса) замість *actor* (актор). Аналогічно, суфікс *-euse*, запозичений із французької мови, позначав жіночу статтю у професіях, як-от: *accoucheuse* (акушерка) замість *accoucheur* (акушер), *chanteuse* (співачка) на противагу *chanter* (співак), або *masseuse* (масажистка) замість *masseur* (масажист). Ще один латинський суфікс, *-trix*, також використовувався для позначення жінок у певних сферах діяльності, наприклад: *aviatrix* (льотчиця) на противагу *aviator* (льотчик) або *administratrix* (адміністраторка) замість *administrator* (адміністратор).

Крім того, гендерна диференціація закріплювалася через вживання складних слів з основою “*lady*” чи “*man*”. Наприклад, терміни на кшталт *handyman* (різнороб), *policeman* (поліціант) або *fireman* (пожежник) відображали уявлення про те, що ці професії були виключно чоловічими; а *cleaning lady* (прибиральниця), *laundry lady* (прачка) – виключно жіночими.

Паралельно із соціокультурними змінами англійськомовна почала еволюціонувати в напрямку гендерної нейтральності. У сучасному суспільстві спостерігається тенденція до відмови від гендерно-маркованих термінів на користь нейтральних варіантів. Наприклад, замість слова *actress* тепер частіше використовують просто *actor*, незалежно від статі людини. Такі терміни як *spokesman* і *fisherman* замінюються відповідно на гендерно-нейтральні аналоги: *spokesperson* та *fisher*. Ця трансформація відображає суспільне прагнення до забезпечення гендерної рівності та подолання стереотипів, пов'язаних зі статтю, у мовній практиці. Таким чином, еволюція англійської мови демонструє адаптацію до соціокультурних змін і спрямованість на формування інклюзивного мовного середовища в сучасному суспільстві.

Для кращого розуміння масштабів і особливостей гендерної сегрегації на ринку праці доцільно звернутися до графічного представлення даних, що демонструє відсоткове співвідношення професій за гендерною ознакою та дозволяє візуалізувати нерівномірний розподіл чоловіків і жінок у різних сферах зайнятості (див. рис. 1). Гендерний розподіл у професійних сферах, представлений на графіку, створено на основі статистичних даних країн Європейського Союзу за 2023 рік (рис. 1).



Рис. 1. Гендерний розподіл у професійних сферах, створений на основі статистичних даних країн Європейського Союзу за 2023 рік

Графік підтверджує висловлений раніше тезис про те, що жінки переважають у сферах, пов'язаних із доглядом і обслуговуванням (медицина, освіта, адміністративні послуги), тоді як чоловіки домінують у технічних, фізично важких і високооплачуваних галузях (будівництво, ІТ, транспорт). Такий дисбаланс підкреслює необхідність подолання гендерної сегрегації на ринку праці шляхом зміни суспільних стереотипів, забезпечення рівного доступу до всіх професій та створення умов для кар'єрного розвитку незалежно від статі.

Таким чином, гендерна сегрегація на ринку праці не лише впливає на професійний вибір жінок і чоловіків, але й знаходить своє відображення у мовних практиках, які закріплюють стереотипні уявлення про роль статей у суспільстві. Отже, у сучасному соціолінгвістичному контексті все ще зберігається схильність до стереотипізації професійних ролей, що виявляється у використанні мовних засобів, які відображають традиційні гендерні уявлення. У підтвердження сказаному вище проведемо експеримент з використанням програми Google Translate (див. рис. 2).

Так, наприклад, в англійській мові слово *"police officer"* є гендерно-нейтральним і може позначати як особу чоловічої, так і жіночої статі, а в українській мові слова *«поліціант»* та *«поліціантка»* використовуються для позначення різних статей представників цієї професії. Це призводить до того, що для виконання перекладу МП обирає довільну стать для позначення представника певної професії, якщо це не було передбачено текстом оригіналу. Такі дослідники як Габріель

Становський, Ноа Сміт та Люк Зеттлемойер назвали це явище «довільним гендерним маркуванням», а такі об'єкти – довільно гендерно-маркованими одиницями [5]. На їхню думку, переклад з довільним гендерним маркуванням є значною проблемою в МП, оскільки збігається зі стереотипами, підкріплюючи шкідливі суспільні упередження [5, с. 1679–1680].

Таке відтворення гендерних стереотипів у мові призводить до неточностей у перекладі, особливо коли йдеться про конкретних осіб. Розглянемо для прикладу переклад речення з англійської мови українською за допомогою Google Translate та DeepL у порівнянні з офіційним перекладом (табл. 1).

У перших двох випадках системи машинного перекладу автоматично приписали чоловічий рід слову *«президент»* та всім дієсловам, хоча доречно було б використовувати жіночий рід, демонструючи, що алгоритми машинного перекладу Google Translate та DeepL все ще базуються на традиційних гендерних стереотипах і не враховують контекст та реальну гендерну приналежність особи, про яку йдеться. Людина-перекладач, на відміну від машинних систем, врахувала, що президентом Грузії є жінка, і використала відповідний фемінітив *«президентка»* та дієслова у жіночому роді [19]. Використання фемінітивів та правильних гендерних форм у перекладі не лише забезпечує точність інформації, але й сприяє подоланню гендерних стереотипів у мові, що є необхідним для офіційних документів та публічних заяв, де точність та коректність формулювань мають першорядне значення.

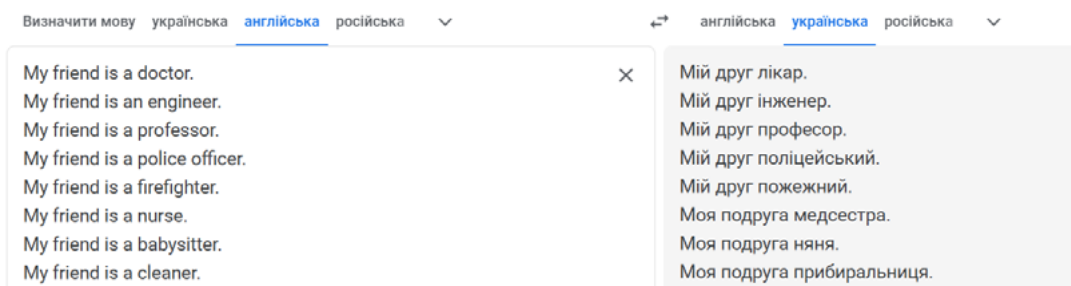


Рис. 2. Приклад гендерно упередженого перекладу англійського тексту українською мовою за допомогою Google Translate

Таблиця 1

Оригінал	The Georgian President <u>thanked</u> for all the support and <u>commended</u> Ukraine's decision to impose sanctions on Bidzina Ivanishvili.
Переклад за допомогою Google Translate	Президент Грузії <u>подякував</u> за підтримку та високо <u>оцінив</u> рішення України запровадити санкції проти Бідзіни Іванішвілі.
Переклад за допомогою DeepL	Президент Грузії <u>подякував</u> за підтримку та <u>привітав</u> рішення України запровадити санкції проти Бідзіни Іванішвілі.
Офіційний переклад з сайту президента України	Президентка Грузії <u>подякувала</u> за всю проявлену підтримку й <u>привітала</u> рішення України застосувати санкції проти Бідзіни Іванішвілі.

Варто зазначити, що системи МП мають низку недоліків, які впливають на точність перекладу, а саме: недостатній рівень врахування контексту, що призводить до помилок у передачі багатозначних слів, ідіом чи метафор [20]; ігнорування культурних особливостей, стилістики та тону тексту оригіналу, через що спотворюється авторська інтенція [21]; труднощі роботи з низькоресурсними мовами, для яких існує обмежена кількість лінгвістичних даних, необхідних для систем обробки природної мови, таких як машинний переклад, розпізнавання мовлення чи аналіз тексту [22, с. 105–106]. Крім того, машинний переклад також не враховує специфічні особливості мов, зокрема використання в українській мові фемінітивів – слів жіночого роду, що альтернативні або парні аналогічним поняттям чоловічого роду (які зазвичай використовуються до всіх людей незалежно від їхньої статі) [23, с. 255]. Основне завдання фемінітивів – зробити жінку та її діяльність видимими у мовному просторі. Вони допомагають подолати стереотипні уявлення про «чоловічість» багатьох посад і професій та підкреслюють рівноправну участь жінок у всіх сферах сучасного життя [16]. Навіть за наявності контексту, який вказує на жіночу стать особи, системи машинного перекладу перекладають назви професій у «стандартному» чоловічому роді за допомогою маскулінітивів – назв осіб чоловічої статі [23, с. 63].

У контексті сучасних тенденцій розвитку української мови, зокрема щодо використання фемінітивів, проведемо порівняльний аналіз перекладу речення “*Oksana Ivanets’ story: a lieutenant colonel, a military journalist, a mother of five*” [24] українською мовою за допомогою чотирьох популярних систем МП (Google Translate, DeepL, Microsoft Translator, OpenL), що має на меті оцінити здатність автоматизованих перекладацьких систем адекватно відтворювати гендерно-чутливу лексику (табл. 2).

Результати дослідження показують, що жодна з проаналізованих систем МП не змогла повністю коректно відтворити фемінітиви у перекладі. Усі чотири системи переклали “*lieutenant colonel*” як «*підполковник*», не використавши фемінітив «*підполковниця*», що вказує на відсутність повної адаптації автоматизованих систем перекладу до відтворення гендерно-чутливої лексики у військовій сфері. Щодо перекладу терміну “*military journalist*”, спостерігаємо певну варіативність. Google Translate та OpenL переклали цей термін як «*військовий журналіст*», тоді як DeepL та Microsoft Translator використали фемінітив «*військова журналістка*», демонструючи, що деякі системи МП вже починають розпізнавати та використовувати фемінітиви у професійній сфері, хоча й не послідовно. Варто зазначити, що всі системи МП коректно переклали “*mother of five*” як «*мати п’ятьох дітей*», за винятком Microsoft Translator, який використав більш розмовний варіант «*мама п’ятьох дітей*». Це демонструє, що системи МП загалом здатні адекватно передавати гендерно-специфічні терміни, коли вони не пов’язані з професійною сферою.

Гендерні упередження у МП виявляються не лише на рівні окремих слів, але й у більш широкому контексті, залежно від того, який обсяг лінгвістичного оточення алгоритм враховує при перекладі конкретного слова [1]. Наприклад, якщо перше речення звучить як “*This is a story about a doctor*”, то як людина-перекладач, так і система МП зіткнуться з неоднозначністю щодо вибору між «*лікар*» і «*лікарка*». Однак, якщо наступне «*уточнювальне*» речення містить додаткову інформацію про стать, наприклад, “*She’s upset because she misdiagnosed a patient*”, людина-перекладач зробить правильний висновок про жіночий рід у першому реченні. Натомість, МП може розглядати кожне речення окремо, залишаючи початкове припущення про рід незмінним. Як показано на рис. 3, іноді об’єднання двох речень в одне може підвищити точність МП.

Таблиця 2

Назва МП	Переклад
Google Translate	Історія Оксани Іванець: <i>підполковник</i> , <i>військовий журналіст</i> , мати п’ятьох дітей.
DeepL	Історія Оксани Іванець: <i>підполковник</i> , <i>військова журналістка</i> , мати п’ятьох дітей.
Microsoft Translator	Історія Оксани Іванець: <i>підполковник</i> , <i>військова журналістка</i> , мама п’ятьох дітей.
OpenL	Історія Оксани Іванець: <i>підполковник</i> , <i>військовий журналіст</i> , мати п’ятьох дітей.

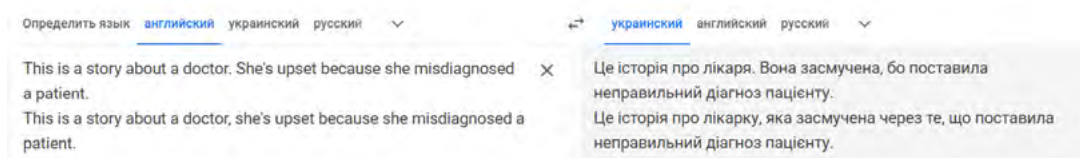


Рис. 3. Приклад перекладу речення українською мовою за допомогою Google Translate

Так, початкове припущення, що лікарем буде чоловік, найімовірніше, спричинене упередженими навчальними даними, які використовуються для тренування алгоритмів з метою виявлення закономірностей, здійснення прогнозів або класифікації інформації [25].

Водночас неспроможність системи машинного перекладу скоригувати це припущення навіть за умови надання додаткового контексту пояснюється її вбудованими налаштуваннями, які орієнтовані на переклад окремих речень, а не тексту в цілому [26].

Висновки. Гендерна упередженість у машинному перекладі формується під впливом двох ключових факторів: специфіки алгоритмів штучного інтелекту та особливостей навчальних даних, що відображають усталені гендерні стереотипи. Це питання особливо гостро виявляється під час перекладу з/на мов з різними граматичними системами, наприклад, з англійської на українську.

Системи машинного перекладу використовують чоловічий рід як стандартний варіант для гендерно-нейтральних слів, що відтворює явище «чоловічого роду за замовчуванням» та призводить до систематичних неточностей, особливо при перекладі назв професій, де технічні та керівні позиції автоматично асоціюються з чоловіками, а професії у сфері догляду чи обслу-

говування – з жінками. Така мовна практика є відображенням глибших соціальних тенденцій, зокрема гендерної сегрегації на ринку праці, яка проявляється у нерівномірному розподілі професій між статями.

Недоліки сучасних систем машинного перекладу охоплюють недостатнє врахування контексту, ігнорування стилістичних і культурних особливостей тексту, а також труднощі роботи з низькоресурсними мовами. Зокрема, українська мова має специфіку використання фемінітивів – слів жіночого роду для позначення професій чи посад, однак автоматизовані системи перекладу часто не враховують ці особливості й використовують маскулінітиви навіть за наявності контексту, що вказує на жіночу стать особи.

Для подолання гендерної упередженості необхідно вдосконалювати моделі машинного перекладу шляхом розширення навчальних корпусів текстами з гендерною інклюзивністю та впровадження адаптивних алгоритмів. Важливим є також аналіз впливу технологій машинного перекладу на формування суспільних уявлень про гендерні ролі. Інтеграція принципів гендерної рівності у мовні технології сприятиме забезпеченню точності, етичності та інклюзивності автоматизованих систем перекладу, що становить перспективи подальших досліджень.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

1. Rescigno A., Monti J. Gender bias in machine translation: A statistical evaluation of Google translate and DeepL for English, Italian and German. *Proceedings of the International Conference on Human-informed Translation and Interpreting Technology 2023*. 2023. P. 1–11.
2. Подвойська О., Козоріз І., Гончаренко Н. Машинний переклад за допомогою інтернет ресурсів. *Молодий вчений*. 2024. № 4 (128). С. 68–71. URL: <https://doi.org/10.32839/2304-5809/2024-4-128-10>.
3. Luccioli A., Xausa C., Dolei E. Investigating Gender Bias in Machine Translation. A Case Study between English and Italian. *Università di Bologna*. 2020. P. 29–49.
4. Губіна О. В. Прагматичні аспекти функціонування англomовних митних термінів. *Актуальні питання іноземної філології*. 2022. С. 57–61. URL: <https://doi.org/10.32782/2410-0927-2021-14-9>.
5. Stanovsky G., Smith N. A., Zettlemoyer L. Evaluating gender bias in machine translation. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. 2019. P. 1679–1684.
6. Liao Y., Callison-Burch P. C., Alur P. R. Gender bias in neural machine translation. *University of Pennsylvania*. 2021. P. 1–41.
7. Triboulet B., Bouillon P. Evaluating the Impact of Stereotypes and Language Combinations on Gender Bias Occurrence in NMT Generic Systems. *Proceedings of the Third Workshop on Language Technology for Equality, Diversity and Inclusion*. 2023. P. 62–70.
8. Ghosh S., Caliskan A. ChatGPT perpetuates gender bias in machine translation and ignores non-gendered pronouns: Findings across Bengali and five other low-resource languages. *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*. 2023. P. 901–912.
9. Saunders D., Byrne B. Reducing gender bias in neural machine translation as a domain adaptation problem. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. 2020. P. 7724–7736.
10. Gete H., Etchegoyhen T. Does context help mitigate gender bias in neural machine translation?. *Findings of the Association for Computational Linguistics: EMNLP*. 2024. P. 14788–14794.
11. Глосарій ґендерних термінів. *Державна судова адміністрація України*. URL: <https://dsa.court.gov.ua/dsa/inshe/gender/glossary>.

12. Гендерні переклади Bing вирішують проблему упередженості в перекладі. *Блог перекладача Microsoft*. URL: <https://www.microsoft.com/uk-ua/translator/blog/2023/03/08/bings-gendered-translations-tackle-bias-in-translation/>.
13. Měchura M. 10 things you should know about gender bias in machine translation. *Fairslator*. URL: <https://www.fairslator.com/10-things-about-gender-bias-in-mt>.
14. Причини виникнення гендерної нерівності у трудових відносинах. *Східне міжрегіональне управління – Державної служби України з питань праці*. URL: <https://smu.dsp.gov.ua/news/prychyny-vynyknennia-hendernoi-nerivnosti-u-trudovykh-vidnosynakh/> (дата звернення: 28.03.2025).
15. Angelko I. V., Lech G. A., Ribun L. V. Гендерна сегрегація – стереотип та проблема сучасного розвитку вітчизняного ринку праці. *Scientific Bulletin of UNFU*. 2018. Т. 28, № 9. С. 31–34. URL: <https://doi.org/10.15421/40280905>.
16. Главацька Н., Подкоритова Л. Сучасні зміни гендерних ролей дорослих. *Psychology Travelogs*. 2021. № 2. С. 25–32. URL: <https://doi.org/10.31891/pt-2021-2-3>.
17. Гендерні ролі та стереотипи. *Охтирська районна державна адміністрація*. URL: <https://oht.sm.gov.ua/index.php/en/gumanitarna-politika/sotsialnyi-zahyst/3350-genderni-rol-i-ta-stereotipi>.
18. Українкам дозволили працювати на "шкідливих" роботах // Українська правда. Життя. 2017. URL: <https://life.pravda.com.ua/society/2017/12/21/228097/>.
19. Володимир Зеленський і Саломе Зурабішвілі обговорили ситуацію в Грузії. *Офіційне інтернет-представництво Президента України*. URL: <https://www.president.gov.ua/news/volodimir-zelenskij-i-salome-zurabishvili-obgovorili-situaci-94889>.
20. Alawi N., Abdulhaq S. Machine Translation: The Cultural and Idiomatic Challenge. *Journal of Al-Azhar University – Gaza (Humanities)*. 2017. Vol. 2, no. 19. P. 1–28. URL: <https://digitalcommons.aaru.edu.jo/alazhar/vol19/iss2/9>.
21. Fedchuk R., Vysotska V. Information Technologies for Solving the Problem of Correcting Errors in Ukrainian-language Texts. *Visnik Nacional'nogo universitetu "L'vivs'ka politehnika". Seriâ Informacijni sistemi ta mereži*. 2024. Vol. 16. P. 11–34. URL: <https://doi.org/10.23939/sisn2024.16.011>.
22. Ухачевич Т. Я., Кустра Н. О. Дослідження впливу обрізання та тонкого налаштування моделі автоматичного розпізнавання мовлення на її точність. *Scientific Bulletin of UNFU*. 2024. Т. 34, № 5. С. 104–109. URL: <https://doi.org/10.36930/40340514>.
23. Шевченко З. В. Словник ґендерних термінів. Черкаси, 2016. 336 с.
24. Підполковниця, військова журналістка та мати 5-х дітей: Історія Оксани Іванець. *Медіацентр Україна*. URL: https://mediacenter.org.ua/uk/?s=Підполковниця,+військова+журналістка+&post_type=pd_themes.
25. Розробка навчального андроїд-застосунку з теми «Сортування вставками» дистанційного навчального курсу «Алгоритми і структури даних» / О. Кошова та ін. *INFORMATION TECHNOLOGY AND SOCIETY*. 2024. № 5 (11). С. 34–42. URL: <https://doi.org/10.32689/maup.it.2023.5.5>.
26. Farkas A., Németh R. How to measure gender bias in machine translation: Real-world oriented machine translators, multiple reference points. *Social Sciences & Humanities Open*. 2022. Vol. 5, no. 1. P. 100239. URL: <https://doi.org/10.1016/j.ssaho.2021.100239>.