



І. А. Редозуб¹, Н. Ф. Хайрова^{2,3}
¹НТУ “ХПІ”, м. Харків, Україна, ivan.red707@gmail.com, ORCID ID: 0009-0008-3312-4700

²НТУ “ХПІ”, м. Харків, Україна;

³Umea University, Sweden, nina.khairova@umu.se, ORCID ID: 0000-0002-9826-0286

ОЦІНЮВАННЯ ЗДАТНОСТІ МОДЕЛЕЙ GENAI ВИЯВЛЯТИ ЕТИЧНІ ВИМІРИ ПУБЛІЧНОГО ДИСКУРСУ: КЕЙС УКРАЇНСЬКИХ БІЖЕНЦІВ

У статті досліджується здатність моделей генеративного штучного інтелекту (GenAI) виявляти й інтерпретувати етичні виміри публічного дискурсу в цифровому середовищі. Актуальність роботи зумовлена зростанням ролі соціальних мереж як джерела суспільних наративів і потребою в аналізі ціннісно навантажених, нормативно чутливих категорій. Зокрема, ми використовуємо модель GPT-3.5-turbo для класифікації текстових даних за визначеними категоріями етичних аспектів міграції. З позицій підходу «етика в дизайні» оцінюється потенціал застосування GenAI для зіставлення громадянських і урядових наративів та аналізу їх узгодженості в контексті підтримки процесів суспільної згуртованості.

ГЕНЕРАТИВНИЙ ШТУЧНИЙ ІНТЕЛЕКТ, ВЕЛИКІ МОВНІ МОДЕЛІ, ЕТИЧНІ ВИМІРИ, АНАЛІЗ СОЦІАЛЬНИХ МЕРЕЖ, ТАКСОНОМІЯ

I. A. Redozub, N. F. Khairova. Evaluating the Capability of GenAI Models to Identify Ethical Dimensions of Public Discourse: The Case of Ukrainian Refugees. This paper investigates the capability of generative artificial intelligence (GenAI) models to identify and interpret ethical dimensions of public discourse in the digital environment. The relevance of the study is driven by the growing role of social media as a source of societal narratives and the need to analyze value-laden and normatively sensitive categories. In particular, we employ the GPT-3.5-turbo model to classify textual data according to predefined categories of ethical aspects of migration. From the perspective of the Ethics-in-Design approach, the study evaluates the potential of GenAI to compare citizen and governmental narratives and to analyze their alignment in the context of supporting processes of social cohesion.

GENERATIVE ARTIFICIAL INTELLIGENCE, LARGE LANGUAGE MODELS, ETHICAL DIMENSIONS, SOCIAL MEDIA ANALYSIS, TAXONOMY

Вступ

В останні роки бурхливий розвиток генеративного штучного інтелекту (Generative artificial intelligence (GenAI)) суттєво розширив інструментарій аналізу даних соціальних мереж. Платформи соціальних медіа щоденно генерують величезні обсяги мультимедіального контенту, що залучає текстові повідомлення, коментарі, зображення та відеоматеріали, які в свою чергу відображають реакції користувачів на соціальні, політичні та культурні події. Аналіз цих даних має значний інтерес для дослідників у галузі соціальних наук, обчислювальної лінгвістики та аналізу даних.

Великі мовні моделі (Large Language Models, LLMs) демонструють здатність відображати те, як люди структурують знання та смисли в межах конкретних предметних областей, втілюючи ці знання через концептуальну організацію мови. Згідно з [1], ця здатність проявляється в умінні моделей відповідати на запитання, генерувати зв'язний текст та виводити інформацію про сутності, об'єкти й властивості, з якими вони безпосередньо не стикалися під час навчання. Хоча мовні моделі не володіють знаннями у позамовному досвіді, логічні структури, які вони уловлюють між лінгвістичними поняттями, можуть наближатися до когнітивної організації концептуальних знань людини. Це свідчить про потенціал використання LLMs не лише як інструментів обробки тексту, а і як засобів моделювання структур смислу в публічному дискурсі.

У цьому зв'язку в останні роки дедалі більше уваги приділяється застосуванню великих мовних моделей для класифікації та відображення абстрактних концептуальних категорій, зокрема поведінкових емоцій, етичної відповідності, соціальних і культурних норм, а також правових оцінок у великих текстових корпусах. Цей новий напрямок досліджень демонструє поступовий зсув від поверхневого мовного моделювання до більш ефективної інтерпретації складних, абстрактних, взаємоперекривних та морально значущих категорій.

Здатність великих мовних моделей працювати з абстрактними та ціннісно навантаженими концептуальними категоріями створює підґрунтя для їх застосування в аналізі соціально значущих дискурсів. Зокрема, LLMs можуть бути використані для інтерпретації суспільних настроїв і позицій, що формуються в онлайн-просторах, де мовні, емоційні та нормативні аспекти тісно переплітаються. Сучасні моделі дають змогу глибше досліджувати суспільні настрої, думки та ставлення до різних соціальних питань, забезпечуючи масштабований і більш контекстно чутливий аналіз порівняно з класичними методами машинного навчання.

Цей перехід від імовірнісного до генеративного моделювання підвищує аналітичну точність, оскільки дає змогу дослідникам виявляти приховані смислові структури та динамічні взаємодії між темами, які

більш адекватно відображають складність публічного дискурсу. Але підхід до аналізу соціально значущих дискурсів відкриває можливості не лише для опису поточних станів суспільних настроїв, а й для виявлення їхньої динаміки в часі та просторових відмінностей між спільнотами. Як показано в роботі [2], застосування генеративних моделей у цьому контексті дозволяє переходити від статичного аналізу окремих текстових фрагментів до системного вивчення еволюції тем, позицій і форм колективного осмислення соціальних проблем у цифровому середовищі. Таким чином, аналітичний потенціал LLMs може бути безпосередньо пов'язаний із ширшими соціальними цілями, зокрема з підтримкою процесів суспільної взаємодії та згуртованості.

Водночас, попри те що наявні дослідження демонструють ефективність підходів GenAI для виявлення структурованих суспільних настроїв, ідентифікації впливових позицій та картографування широкого публічного дискурсу, вони здебільшого не зосереджуються на тому, яким чином такі інструменти можуть бути використані для систематичного зіставлення громадянських наративів із конкретними державними нормами крізь призму етичних, правових, моральних і соціальних вимірів, а також для оцінювання потенціалу таких зіставлень у контексті зміцнення суспільної згуртованості.

Особливої актуальності ці обмеження набувають у випадках, коли предметом аналізу є публічний дискурс маргіналізованих спільнот, зокрема груп мігрантів, де питання представлення інтересів і суспільної згуртованості є критично важливими. Ці труднощі зумовлені не лише обмеженою доступністю та видимістю власних позицій мігрантів у публічному дискурсі, а й складністю інтеграції таких даних з офіційними наборами даних.

У відповідь на ці методологічні виклики в даному дослідженні виконується перевірка здатності моделей GenAI виявляти та інтерпретувати етичні виміри в текстовому контенті. При цьому емпірична перевірка цього потенціалу здійснюється на матеріалі публічних і офіційних наративів, пов'язаних із проблематикою вимушеної міграції, зокрема українських біженців, які були змушені залишити країну внаслідок повномасштабного російського вторгнення.

Випадок українських біженців становить репрезентативний і водночас методологічно релевантний кейс з огляду на масштаб і терміновість гуманітарної кризи: за даними Управління Верховного комісара ООН у справах біженців (УВКБ ООН, 2024), до кінця 2024 року понад шість мільйонів українців зазнали вимушеного переміщення. Такий контекст формує насичене та різноманітне середовище даних для перевірки здатності моделей GenAI працювати з нормативно навантаженими та ціннісно чутливими категоріями.

З позицій підходу «етика в дизайні» (Ethics-in-Design) дослідження спрямоване на оцінювання того, наскільки інструменти штучного інтелекту можуть бути використані для аналізу соціально значущих проблем [3] і, ширше, для підтримки процесів суспільної згуртованості шляхом виявлення потенційних розбіжностей між офіційними нормативними рамками та наративами маргіналізованих груп.

1. Аналіз наукових підходів до проблеми

У сучасних дослідженнях дедалі більше уваги приділяється питанням етики застосування штучного інтелекту в гуманітарній сфері та контексті міграції. Зокрема, у статті Міжнародного Червоного Хреста [4] наголошується, що у кризових ситуаціях використання алгоритмів повинно ґрунтуватися на нормах міжнародного права з прав людини, а також доповнюватися етичними принципами прозорості, підзвітності та справедливості. Подібні рекомендації містяться й у матеріалі ASAT Labs — незалежної дослідницько-аналітичної ініціативи у сфері етичного використання штучного інтелекту в гуманітарних проєктах. В цих матеріалах особлива увага приділяється людиноцентризму дизайну, захисту даних біженців та інклюзивності під час проєктування систем [5]. Приклад практичного впровадження таких підходів можна знайти у ініціативах Міжнародного комітету порятунку (IRC), який застосував чатботи у програмі Signpost для підтримки переміщених осіб, проте водночас визначив необхідність анонімізації даних та залучення людини до перевірки відповідей, аби уникнути ризиків поширення неправдивої або шкідливої інформації [6].

Дослідники також звертають увагу на проблему репрезентації біженців у цифровому середовищі. Так, у блозі Університету Організації Об'єднаних Націй (United Nations University, UNU) зазначається, що створення AI-агентів чи цифрових аватарів біженців може призвести до їх відчуження від власного голосу та посилення стереотипів, якщо такі технології не контролюються самими постраждалими спільнотами [7].

У роботі [8] продемонстровано, як загальні принципи етики штучного інтелекту (автономія, справедливість, запобігання шкоді, прозорість) можна трансформувати у конкретні технічні вимоги для систем прогнозування у сфері міграційного менеджменту. Це відповідає ширшій тенденції до розвитку підходів Value Sensitive Design, спрямованих на вбудовування людських цінностей у процес проєктування технологій [9]. Подальші емпіричні дослідження засвідчують, що для практиків і законодавців ключовими проблемними аспектами залишаються прозорість, підзвітність і захист конфіденційності, тоді як одним із найбільших викликів є брак ефективних механізмів регуляторного забезпечення [10].

Важливим для розуміння загальної картини є також Звіт про стан етики штучного інтелекту Монреальського інституту етики ШІ, який відстежує ключові тенденції в екосистемі, зокрема питання упевреності, регуляторного середовища та суспільного сприйняття. [11].

У ряді останніх досліджень значну увагу також приділено етичним рамкам і практикам, що стосуються мігрантів та біженців, особливо молоді та вразливих груп. Наприклад, стаття [12] аналізує, як методологічні підходи до роботи з молоддю-біженцями можуть враховувати етичні дилеми, такі як асиметрія влади, питання відповідальності, приватності й довірчих відносин між дослідниками та учасниками. Інше дослідження [13] звертається до інклюзивності та етичних норм у контексті складної доступності мігрантів під час гуманітарних криз, підкреслюючи необхідність забезпечення захисту прав, гідності та врахування голосів тих, хто зазвичай маргіналізується.

Також важливо звернути увагу на праці, що розглядають етичні аспекти застосування ШІ в міграційному менеджменті. Зокрема, стаття [8] обговорює, як принципи етики, так як прозорість, справедливість, недопущення шкоди можуть бути вбудовані у предикативні інструменти, які використовуються гуманітарними організаціями, щоб мінімізувати ризики і захистити права мігрантів.

Крім того, дослідження [14] ілюструє, як медіа через візуальні образи відображають біженців, надаючи їм різні суб'єктивні позиції (вразливі жертви, особи, що зазнають страждань, резиліентні), що також має етичне значення у формуванні суспільного сприйняття.

У локальному українському контексті прикметним є проведення Літньої школи «Ethics&AI Summer School»¹ в Українському католицькому університеті у 2023 році, де обговорювалися питання довіри, корупційних ризиків та синтетичного контенту. Крім того, у 2025 році Українська національна асоціація адвокатів приєдналася до розробки глобальних стандартів етики використання штучного інтелекту спільно з міжнародними організаціями².

2. Методологія дослідження

Дослідження проходило в кілька етапів: спочатку було визначено географічні межі роботи, далі - знайдено та сформовано перелік офіційних джерел інформації й Telegram-каналів та чатів, після чого здійснено збір та обробку даних. Схематично повна методологія відображена на рис. 1.

Створення датасету передусім починається з визначення напрямків міграції біженців, що є надзвичайно важливим для перевірки точності даних і достовірності інформації. У межах дослідження були визначені основні напрями міграції українців у Європі та Північній Америці, зокрема: Україна, Польща, Канада, Сполучені Штати Америки, Словаччина, Молдова, Німеччина, Чехія, Іспанія, Велика Британія, Італія, Швеція та Ірландія³. Отримані результати

- 1 UCU Ethics&AI Summer School 2023 - <https://apps.ucu.edu.ua/events/ucu-ethics-ai-summer-school-2023/>
- 2 Ukrainian advocates join efforts to develop AI ethics standards – <https://en.unba.org.ua/activity/news/10435-ukrainian-advocates-join-efforts-to-develop-ai-ethics-standards.html>
- 3 Eurostat – https://ec.europa.eu/eurostat/databrowser/view/MIGR_ASYTPSM__custom_7394287/default/table?lang=en

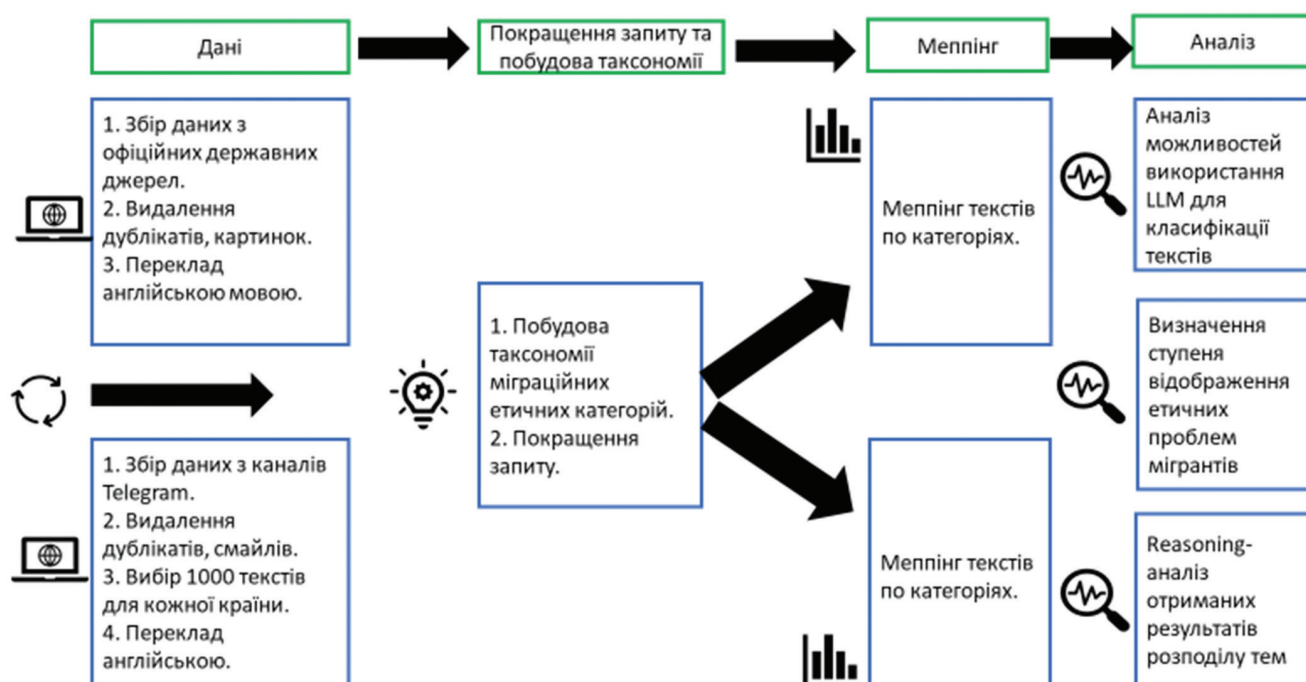


Рис. 1. Чотири етапи методології дослідження

також створюють базу для подальшого порівняльного аналізу між офіційними джерелами різних країн і реальними тенденціями, відображеними у відкритих комунікаційних середовищах, таких як соціальні мережі та месенджери.

Для початку ми вирішили створити 2 датасети, де перший містить публічну інформацію від медіаресурсів держав та міжнародної організації з міграції [15], а другий включає дискурс мігрантів у соціальній мережі Telegram [16].

Для першого датасету ми обробили більше 30 джерел таких як:

- офіційні вебсайти міграційних служб і міністерств закордонних справ різних країн;
- посольства України в країнах перебування;
- YouTube-канали державних установ;
- акаунти урядовців у мережі Twitter;
- державні вебресурси, зокрема сайти міністерств закордонних справ, охорони здоров'я та освіти;
- українські урядові структури, що надають підтримку біженцям за кордоном;
- вебсайт Міжнародної організації з міграції (ІОМ), включно з її українським представництвом.

В результаті було отримано 708 текстів, з яких близько третини припадає на статті з сайту ІОМ.

Формати проаналізованих даних включали тексти вебсайтів, PDF-файли, твіти, а також аудіо- та відеофайли з YouTube. Збирання текстів із вебсайтів здійснювалося за допомогою бібліотеки Python BeautifulSoup. Для обробки PDF-документів використовували бібліотеку PyPDF2, тоді як обробку сирого тексту забезпечувала бібліотека WordNinja, а визначення мови контенту виконувала бібліотека LangDetect. Текстові дані з відеофайлів отримували через вебсервіс Sonix, який надає транскрипцію аудіо у текст на кількох мовах. Для невеликих джерел або вебсайтів із обмеженнями на скрапінг застосовувався ручний парсинг. Кроки попередньої обробки тексту включали очищення (видалення емодзі, символів та іконок), лематизацію та видалення порожніх і дубльованих записів.

Для датасету публічного дискурсу було зібрано великий масив повідомлень із Telegram-груп, де українські біженці обмінюються інформацією та обговорюють актуальні питання. Такі чати стали основою спілкування і взаємної допомоги біженців у приймаючих країнах. Для пошуку відповідних каналів використовували ключові слова. Після початку повномасштабного вторгнення Росії активність у таких групах значно зросла, що перетворило Telegram на важливий канал комунікації для українців [17]. Було зібрано 632 372 повідомлення в соціальній мережі. Для забезпечення рівномірного розподілу текстів по країнах було зроблено обмеження до 1000 випадкових текстів з кожної з країн, що призвело до створення датасету в 12000 повідомлень. Табл. 1 відображає розмір остаточно використаних у дослідженні датасетів.

Таблиця 1

Об'єм створених датасетів

Назва датасету	Кількість текстів
Датасет офіційних джерел	708
Датасет соціальної мережі	12000

Для вибору мовної моделі було здійснено порівняльний аналіз затримки відповіді (latency) та вартості використання популярних великих мовних моделей, що застосовуються для автоматизованого аналізу текстів. Оцінювання цих параметрів є критично важливим для досліджень, орієнтованих на обробку корпусів текстових даних.

З погляду затримки відповіді, GPT-3.5 Turbo демонструє одну з найнижчих затримок серед великих мовних моделей в відкритих бенчмарках, що пояснюється його оптимізацією для швидкої генерації відповідей. Модель Claude 3 також є конкурентною за швидкістю, однак фактичні показники суттєво залежать від конкретного підтипу (Opus або Sonnet) та платформи розгортання. Для LLaMA-2 швидкодія значною мірою визначається розміром моделі та апаратним забезпеченням: компактні версії (7B) можуть забезпечувати дуже низьку затримку на оптимізованих GPU, тоді як великі моделі часто мають більшу затримку відповіді без пропорційного зростання точності.

Не менш важливим є економічний аспект. GPT-3.5 Turbo має фіксовану ціну \$0.50 за 1 млн вхідних токенів та \$1.50 за 1 млн вихідних токенів, тоді як Claude 3 є значно дорожчим і коштує \$3 та \$15 відповідно. Для LLaMA-2 централізоване ціноутворення відсутнє, а вартість залежить від провайдера. Найнижча зафіксована ціна становила \$0.36 за 1 млн токенів як для входу, так і для виходу. За умовного сценарію аналізу тексту з 250 вхідними та 50 вихідними токенами орієнтовна вартість одного запиту становить \$0.0002 для GPT-3.5 Turbo, \$0.0015 для Claude 3 та \$0.000108 для LLaMA-2. У табл. 2 представлено порівняння трьох великих мовних моделей за показниками, релевантними для застосування запропонованої методики.

З огляду на наведене порівняння, у межах даного дослідження було обрано модель GPT-3.5 Turbo як основний інструмент аналізу. Такий вибір зумовлений оптимальним поєднанням низької затримки відповіді, стабільної якості результатів та високої економічної ефективності, а також легкості використання при масовій обробці текстів. На відміну від дорожчих моделей, ChatGPT забезпечує достатній рівень точності для задач тематичної класифікації на основі таксономії без суттєвого зростання обчислювальних витрат. Порівняно з self-hosted рішеннями на базі LLaMA-2, використання ChatGPT дозволяє уникнути додаткових витрат на інфраструктуру та технічне обслуговування, що є критично важливим для відтворюваності та масштабованості дослідження.

Таблиця 2

Порівняння великих мовних моделей

Характеристика	GPT-3.5	Claude 3	LLaMA-2
Легкість використання	Найзручніший API на ринку: SDK для більшості мов, просте налаштування, детальна документація.	Повноцінний API від Anthropic. Простий REST/SDK, гарна документація, потокове передавання, стабільний час безперебійної роботи.	Немає офіційного централізованого API. Тільки самостійний хостинг або через провайдерів. Зручність залежить від використання та потребує глибшого дослідження.
Швидкість	Найнижча затримка серед масових моделей, стабільна швидкість навіть під навантаженням.	Високий, але трохи повільніший, ніж GPT-3.5 Turbo	Сильно залежить від хостингу та розміру моделі: 7B – швидкий, 70B – повільніший
Ціна за 1 млн токенів	\$3 вхід/\$15 вихід	\$0.5 вхід/\$1.5 вихід	Від \$0.36 за вихід та вхід
Приблизна вартість аналізу (250 вхід + 50 вихід токенів)	\$0.0002 (0.02 цента)	\$0.0015 (0.15 цента)	\$0.000108 (0.0108 цента)

Таким чином, для визначення етичної складової текстів ми використали OpenAI API та модель “GPT-3.5-turbo-16k”. Ця модель була обрана через її здатність обробляти великі тексти, що було необхідно для визначення тематики довгих документів. Крім того, модель оптимізована для швидшої обробки, що дозволяє ефективно класифікувати тексти за задалегідь визначеними категоріями. При цьому вона зберігає весь спектр можливостей, включно з розумінням мови, генерацією тексту, узагальненням та іншими функціями.

На другому етапі нами було створено таксономію тем етичного дискурсу в текстах, пов’язаних з міграцією. Для цього ми виокремили десять наукових статей, які були визначені академічною пошуковою платформою Google Scholar за період з 2022 по 2025, що стосуються етичних проблем біженців. Після цього з використанням GenAI за допомогою промптів ми виокремили більше 200 сутностей з цих статей. Деякі з них містили дублікати та сутності, які не стосувалися міграції чи етики. Після видалення зайвих сутностей ми отримали більше 66 сутностей, які були розподілені на 7 типів. Табл. 3 вказує на верхній рівень класифікації сутностей, систематично організованих навколо центральної теми міграції та чітко розподілених за етичними аспектами.

Таблиця 3

Кількість сутностей таксономії

Назва типу	Кількість сутностей
Міграційний статус та наміри	10
Етичні міркування	19
Міграція та емоційні наслідки	8
Етика досліджень	13
Міграційна політика	4
Права людини та соціальна справедливість	8
Методологія дослідження міграції	4
Всього	66

На наступному етапі нашого дослідження ми проаналізували ключові типи, відібрані системою обробки природної мови, представленою моделлю GPT. Оскільки ми прагнули застосовувати механізми штучного інтелекту для підтримки урядових структур, ми додали експертну перевірку правильності інтерпретації концептів, пов’язаних з етикою міграції [18].

1. *Міграційний статус та наміри.* Цей тип охоплює різні статуси міграції, такі як біженці, шукачі притулку, переміщені особи, особи з нестабільним статусом, а також добровільну міграцію, включно з економічними мігрантами, возз’єднанням сімей та процесами інтеграції культурної та економічної. Він дозволяє зрозуміти причини, мотиви та правовий статус мігрантів у новому середовищі.

2. *Етичні міркування.* Цей тип зосереджується на етичних принципах та практиках, включно з етичною відповідальністю, дилемами, викликами та стратегіями пом’якшення. Він також охоплює етичні теорії, такі як догляд, автономія, справедливість, солідарність та довіра, що формують основу для морального аналізу міграційних процесів.

3. *Міграція та емоційні наслідки.* Тут розглядаються емоційні впливи міграції, зокрема емоційні зв’язки, психологічні труднощі, культурний шок, стрес, конфлікти, відчуження та адаптація до повсякденного життя, особливо для малодохідних груп населення. Це допомагає оцінити психосоціальний стан мігрантів і їхню здатність інтегруватися в нове середовище.

4. *Етика досліджень.* Цей тип включає етичні міркування у дослідженнях міграції, такі як інформована добровільна згода, конфіденційність, приватність, оцінка ризиків та переваг, компетентність та асиметрії влади. Також розглядаються специфічні етичні аспекти міграції, наприклад етична трансформація, обмін даними, маскуваність даних та прозорість.

5. *Міграційна політика.* Цей тип аналізує вплив міграційної політики, включно з розробкою політик, їхнім впливом, рекомендаціями та міграційними нарративами. Він дозволяє зрозуміти, як політичні рішення формують умови для мігрантів та відображають соціальні пріоритети.

6. *Права людини та соціальна справедливість.* Тут розглядаються питання вразливих груп населення, включно з особами без документів, вразливими групами та громадами прийому, а також проблеми соціальної справедливості, такі як рівність можливостей, культурне різноманіття та суспільні уявлення. Це

підкреслює важливість захисту прав мігрантів і забезпечення рівності в суспільстві.

7. *Методології дослідження міграції.* Цей тип охоплює методи дослідження міграції, включно з моніторингом міграційних потоків, використанням політичних інформаторів, міжурядовим збором даних та статистикою громадянського суспільства. Він спрямований на забезпечення об'єктивного аналізу міграційних процесів для прийняття обґрунтованих рішень.

Ієрархія даної таксономії, що включає в себе 4 рівні, зображена на рис. 2:

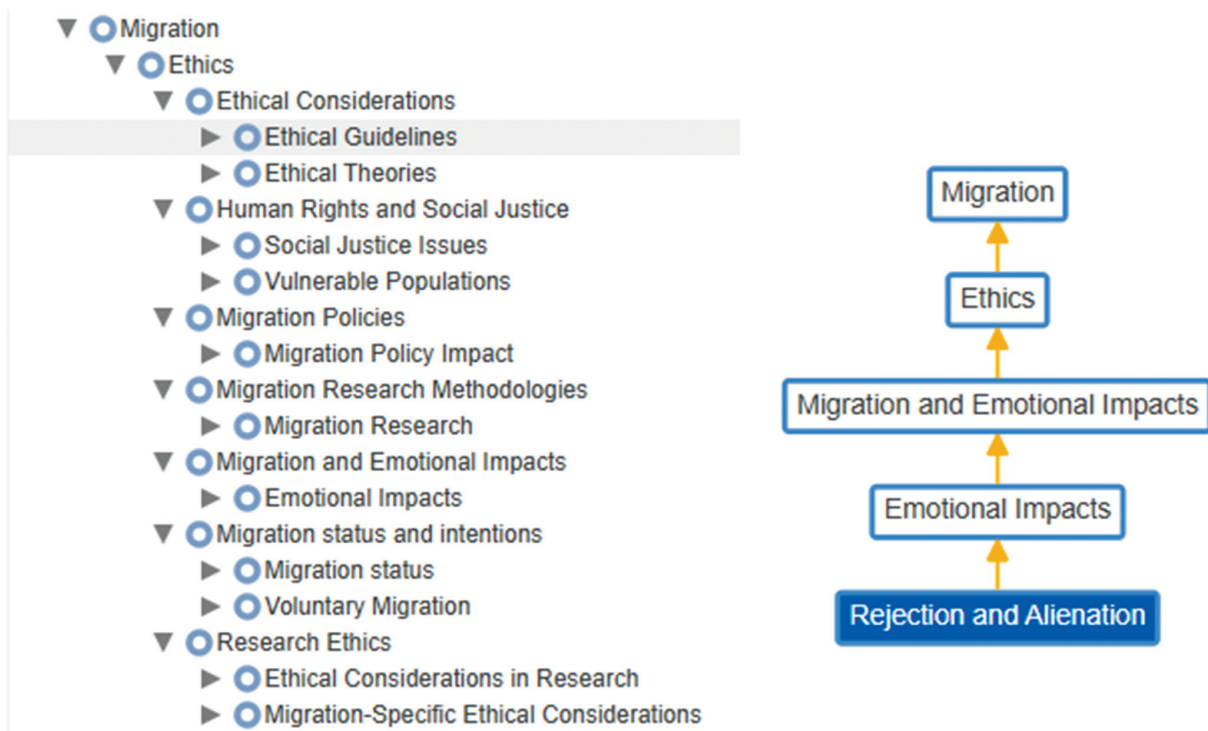


Рис. 2. Ієрархія розробленої таксономії

Наступним кроком було визначено, які тексти з двох датасетів мають відношення до етичних аспектів, а які відносяться до інших питань міграції.

Взаємодія з моделлю GPT організована за допомогою основної функції для визначення типу етичного контексту тексту. Для цього було використано наступний промпт :

“Here is a taxonomy of types of ethics topics in migration texts. Taxonomy:...

I will send you a file with texts and your task is to determine which type every text is and provide reasoning for choosing this type. Types are under numbers 1-7. Don't choose any other options or return None for the type or reason. Always choose between 7 types. Provide me with a CSV file with columns: text_number, text, type_name, reason”.

Створена функція також враховувала можливі обмеження по токенах та дотримувалася правил

використання API. Цінними є також отримані інсайти щодо причин, з яких модель обирає конкретний тип.

Ми провели кількісний аналіз того, як часто публікації офіційних органів влади та дискурси користувачів соціальних мереж спиралися на нарративи етики в кожному окремому наборі даних. Такий підхід дозволив оцінити частоту обговорення етичних аспектів міграції у різних типах дискурсів та визначити, наскільки добре сприймається відображення етичних проблем у суспільстві.

Додатково ми здійснили якісний аналіз текстів у цих двох наборах даних, щоб ідентифікувати, які складові елементи етики зустрічаються у публікаціях влади та обговореннях у соціальних мережах у нарративах міграції. Це допомогло структурувати матеріал та підготувати його для подальшого порівняльного аналізу. Цей підхід дозволив виявити не лише частоту згадок, а й якісну різницю у фокусі обговорення

етичних питань між офіційними та громадськими дискурсами. Завдяки здатності моделі GPT до аргументування та міркування, ми отримали можливість враховувати контекст, взаємозв'язки між поняттями та тональність висловлювань, що забезпечує більш точний та комплексний аналіз текстового матеріалу.

3. Результати

В результаті дослідження було анотовано 2 тексти з офіційних джерел та 37 текстів з повідомлень у соціальній мережі. Що зрозуміло, свідчить про дуже низький рівень висвітлення етичних проблем як в дискурсі політиків, так і в дискурсі мігрантів. Табл. 4 висвітлює отриманий результат з відображення етичних елементів текстів, які пов'язані з міграцією.

Таблиця 4

Рівень висвітлення етичних аспектів у суспільстві

Назва датасету	Кількість текстів на тему етики	Загальна кількість текстів	% етичних текстів
Датасет офіційних джерел	2	708	0.27%
Датасет соціальної мережі	37	12000	0.31%

Щодо етичних тем що були відображені у цих текстах, то в офіційних джерел по одному тексту було з тем “Міграційний статус та наміри” та “Міграційна політика”. Розподіл тем, що пов'язані з етичною складовою у чатах неофіційних повідомлень Telegram відображено у табл. 5.

Таблиця 5

Розподіл текстів за темами етичного дискурсу

Тема етики	% теми серед текстів, що мають етичну складову
Етичні міркування	70.5%
Міграція та емоційні наслідки	11%
Міграційна політика	8%
Права людини та соціальна справедливість	5.5%
Міграційний статус та наміри	2.5%
Етика досліджень	2.5%
Методологія дослідження міграції	0%

Таким чином, найбільшу кількість текстів система віднесла до загальних етичних та комунікативних орієнтирів, тоді як меншою мірою були представлені теми, пов'язані з емоційним досвідом, міграційною політикою та соціальною справедливістю. Окремі тексти торкалися питань статусу мігрантів, інтеграції та дослідницької етики.

Додатково слід зазначити, що аргументування у межах етичної проблематики міграції є складним,

оскільки значна частина текстів апелює до абстрактних моральних орієнтирів, які не мають чітких меж між категоріями. Такі аргументи, що стосуються справедливості та вразливих груп, часто перетинаються з аргументами, що розглядають етичні та соціальні питання. Аналогічно, аргументи, що стосуються загальних моральних, комунікативних або поведінкових орієнтирів, можуть одночасно включати елементи емоційних переживань, адаптації або психологічні аспекти, що ускладнює їх однозначну інтерпретацію та класифікацію. Табл. 6 відображає аргументування вибору теми, що створено GenAI, з використанням експертної складової

Таблиця 6

Аргументація вибору теми

Аргументування	% використання
Загальні моральні, комунікативні або поведінкові орієнтири	59.5%
Розглядає етичні, моральні або питання соціальної справедливості	10.8%
Описує емоційні переживання, адаптацію або психологічні аспекти	10.8%
Згадує міграційну політику, безпечно працевлаштування або інституційні рамки	8.1%
Стосується справедливості, рівності або вразливих груп населення	5.4%
Обговорює міграційний статус, питання біженців або інтеграційні аспекти	2.7%
Розглядає дослідницькі практики, інформовану згоду або етику даних	2.7%

Висновки

Використання моделі GPT для аналізу текстів дозволило застосувати підхід, заснований на міркуванні та семантичному виведенні, який виходить за межі простого підрахунку слів і зосереджується на смислових зв'язках та контексті етичних понять із використанням таксономії сутностей у межах етичного аспекту міграції. Завдяки цьому стало можливим виявити, як різні типи дискурсу — офіційний і публічний — репрезентують моральні аспекти міграції та які теми при цьому актуалізуються. Такий підхід не лише поглибив розуміння структури комунікації, а й відкрив перспективи для використання штучного інтелекту як аналітичного інструмента в гуманітарних дослідженнях.

Використання розробленої таксономії відіграє ключову роль у підвищенні точності роботи моделі, орієнтованої на міркування та семантичне виведення, оскільки вона задає чіткі концептуальні межі для класифікації етичних тем. Таксономія допомагає моделі не лише точніше розпізнавати релевантні контексти, а й формувати більш послідовну логіку аргументації під час віднесення текстів до певних категорій. Крім того,

така структурована система полегшує інтерпретацію результатів як для дослідників, так і для потенційних державних застосувань, де важливо забезпечити прозорість та відтворюваність процесів автоматизованого аналізу етичних аспектів у публічному дискурсі.

Крім того, проведене дослідження дозволило виявити особливості того, як етичні питання міграції біженців, спричинені війною, висвітлюються у двох вимірах суспільної комунікації. Аналіз показав, що в обох типах джерел етична тематика трактується мало, а її присутність часто залишається опосередкованою через згадки про справедливість, солідарність, турботу чи соціальну підтримку. Це свідчить про потребу у більш системному врахуванні морального виміру міграційної політики, особливо в умовах гуманітарних криз.

Порівняльний аналіз між публічним та офіційним дискурсами висвітлює значну асиметрію у розподілі етичних тем. У соціальних мережах переважає категорія “Етичні міркування”, що свідчить про підвищений інтерес користувачів до моральних аспектів міграції, зокрема до етичних дилем, відповідальності та справедливості. Меншою мірою обговорювалися теми емоційних наслідків міграції, міграційної політики та прав людини й соціальної справедливості. Натомість у офіційному дискурсі спостерігається обмежене представлення етичних питань: виявлено лише два тексти. Один, що стосується міграційного статусу та намірів, та другий, що стосується міграційної політики. Такий дисбаланс підкреслює, що етичний вимір більше зосереджений на етичних міркуваннях у громадському середовищі, ніж у комунікаціях державних структур, що може вказувати на потребу посилення етичної складової в офіційному обговоренні міграційних процесів.

Результати дослідження можуть бути використані державними інституціями для покращення публічної комунікації щодо міграційної політики, зокрема шляхом формування більш чутливих до етичних контекстів повідомлень. Для суспільства ці результати є важливими у підвищенні рівня усвідомлення моральних аспектів прийняття біженців, формуванні толерантності та розвитку культури солідарності. Такі знання також можуть сприяти вдосконаленню механізмів взаємодії між владою, громадськістю та міжнародними організаціями у питаннях гуманітарної підтримки.

Разом із тим, узагальнення та використання отриманих результатів у практиці державного управління та суспільної комунікації потребують обережності з огляду на низку методологічних і емпіричних обмежень дослідження.

По-перше, емпіричний аналіз був зосереджений виключно на етичних вимірах публічного дискурсу, тоді як правові та соціальні аспекти розглядалися переважно на концептуальному рівні. Такий фокус

звужує можливості комплексного аналізу нормативних взаємозв'язків і потребує розширення в подальших дослідженнях.

По-друге, операціоналізація абстрактних етичних категорій залишається методологічно складним завданням. Формалізація ціннісно навантажених понять у вигляді таксономії є проблематичною не лише для моделей штучного інтелекту, а й для експертів-людей, що може впливати на стабільність класифікації та інтерпретованість результатів. Додатково, окремі категорії мають перехресний характер, що зумовлює неоднозначності під час маркування та потребує подальшого уточнення таксономії й розвитку підходів «людина—ШІ».

По-третє, узагальнюваність результатів обмежується репрезентативністю та доступністю даних соціальних мереж. Аудиторії платформ не є демографічно нейтральними, а доступ до агрегованих даних часто обмежується великими технологічними компаніями, що може призводити до часткового або упередженого відображення публічного дискурсу [19].

Крім того, багатомовність корпусу (українська/російська та англійська мови) могла спричинити семантичні розбіжності, не повністю усунуті на етапі попередньої обробки, зокрема під час перекладу нормативно навантажених висловлювань.

Нарешті, попри залучення матеріалів із кількох країн за період 2022–2024 років, дослідження обмежене часовими рамками та не враховує ширших культурних і політичних чинників, що можуть впливати на зіставлюваність наративів між різними національними контекстами.

Подальші роботи можуть розширити часовий горизонт, застосувати багатомовні моделі та глибше інтегрувати контекстуальні змінні для підвищення валідності та відтворюваності результатів.

Отже, дослідження продемонструвало, що поєднання штучного інтелекту та етичного аналізу створює нові можливості для більш глибокого розуміння суспільних процесів, пов'язаних із міграцією. Врахування етичних вимірів у політичних рішеннях та публічних комунікаціях може стати важливою передумовою підвищення соціальної згуртованості та зміцнення довіри між державою та громадянами.

Список літератури:

- [1] Du, C., Fu, K., Wen, B., Sun, Y., Peng, J., Wei, W., Gao, Y., Wang, S., Zhang, C., Li, J. and Qiu, S., (2024), “Human-like object concept representations emerge naturally in multimodal large language models”. arXiv preprint arXiv:2407.01067.
- [2] Khairova, N., Rizun, N., Alexopoulos, C., Ciesielska, M., Lukashchuk, A. and Redozub, I., (2024), “Understanding the Ukrainian Migrants Challenges in the EU: A Topic Modeling Approach”, In Proceedings of the 25th Annual International Conference on Digital Government Research (pp. 196-205). <https://dl.acm.org/doi/10.1145/3657054.3657252>

- [3] Dignum, V. (2018). Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20(1), 1-3.
- [4] International Committee of the Red Cross. (2020). AI for humanitarian action: Human rights and ethics. *International Review of the Red Cross*. <https://international-review.icrc.org/articles/ai-humanitarian-action-human-rights-ethics-913>
- [5] ASAT Labs. (2021). 5 Ethical AI Principles Every NGO Should Follow in Humanitarian Projects. <https://asatlabs.org/5-ethical-ai-principles-every-ngo-should-follow-in-humanitarian-projects>
- [6] AP News. (2023). Refugee aid groups turn to AI chatbots, raising ethical questions. <https://apnews.com/article/9c5fae949c429c38019feabaec5fafa>
- [7] United Nations University. (2022). The role of AI agents in humanitarian action: Balancing ethics and innovation. <https://unu.edu/cpr/blog-post/role-ai-agents-humanitarian-action-balancing-ethics-and-innovation>
- [8] Guillén, M., & Teodoro, A. (2023). Embedding Ethical Principles into AI Predictive Tools for Migration Management in Humanitarian Action. *Social Sciences*, 12(2), 53. <https://doi.org/10.3390/socsci12020053>
- [9] Friedman, B., & Hendry, D. G. (2019). *Value Sensitive Design: Shaping Technology with Moral Imagination*. MIT Press.
- [10] Jobin, A., Ienca, M., & Vayena, E. (2022). AI Ethics: An Empirical Study on the Views of Practitioners and Lawmakers. arXiv preprint arXiv:2207.01493. <https://arxiv.org/abs/2207.01493>
- [11] Montreal AI Ethics Institute. (2022). State of AI Ethics Report (Volume 6). <https://arxiv.org/abs/2202.07435>
- [12] Huizinga R. P. et al. Ethical and Methodological Considerations in Research with Asylum-Seeking and Refugee Youth in European Cities // *Social Sciences*. – 2025. – Т. 14. – № 4. – С. 204.
- [13] Ullah A. K. M. A. Ethical framework and inclusivity: research mechanics of difficult-to-reach migrants in civil military context // *Journal of International Humanitarian Action*. – 2024. – Т. 9. – №. 1. – С. 11.
- [14] Martikainen J., Sakki I. (2024) Visual humanization of refugees: A visual rhetorical analysis of media discourse on the war in Ukraine // *British Journal of Social Psychology*. – 2024. – Т. 63. – №. 1. – С. 106-130.
- [15] Redozub, I., and Khairova, N. (2025), “Ukrainian Migrants: Official Sources Discourse (2022–2023)”, Zenodo, available at: <https://doi.org/10.5281/zenodo.17565879> (accessed 9 November, 2025)
- [16] Khairova, N., Lukashchuk, A., & Redozub, I., (2025), “Ukrainian migrants Telegram discourse (2022-2023)”, Zenodo, available at: <https://doi.org/10.5281/zenodo.17566020> (accessed 9 November, 2025)
- [17] Ghasiya, P. and Sasahara., K. (2023) Messaging strategies of ukraine and russia on telegram during the 2022 russian invasion of ukraine. *First Monday*, 2023.
- [18] Dignum, V., Baldoni, M., Baroglio, C., Caon, M., Chatila, R., Dennis, L., ... & De Wildt, T. (2018, December). Ethics by design: Necessity or curse?. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 60-66).
- [19] Lutz, C., (2022), “Inequalities in social media use and their implications for digital methods research”, in A. Quan-Haase & L. Sloan (Eds.), *The SAGE Handbook of Social Media Research Methods*, pp. 679-690, SAGE.

Received (Надійшла) 28.01.2026

Accepted for publication (Прийнята до друку) 26.02.2026

Publication date (Дата публікації) 27.03.2026