УДК 070:81'42:004.89
DOI https://doi.org/10.32782/2710-4656/2025.5.2/46

Тартачний О. В.
Національний університет «Києво-Могилянська академія»

ВИКРИТТЯ НАРАТИВІВ ТА АВТОМАТИЗОВАНА ОБРОБКА ТЕКСТОВОГО КОНТЕНТУ ЗА ДОПОМОГОЮ ШІ: МОЖЛИВОСТІ ЗАСТОСУВАННЯ ТА СТАНОВЛЕННЯ РЕДАКЦІЙНОЇ ПРАКТИКИ (ЗА МАТЕРІАЛАМИ TEXTY.ORG.UA)

Стаття пропонує практичне осмислення використання великих мовних моделей (large language model, LLM) у журналістиці для автоматизованої обробки текстів і виявлення та картографування маніпулятивних наративів.

Вихідна проблематика полягає в тому, що більшість редакцій інтегрують ШІ на рівні транскрибування, перекладів чи технічної вичитки, але рідко застосовують його як інструмент пошуку тем чи інтерпретації даних. У фокусі дослідження – досвід української редакції, яка системно поєднує можливості великих мовних моделей із редакторською експертизою. На прикладах Texty.org.ua показано, що великі мовні моделі здатні розпізнавати поширені техніки маніпуляції навіть у коротких повідомленнях соціальних платформ.

У статті поєднано аналітичний огляд релевантних підходів іноземних редакцій із поглибленим емпіричним розглядом принципів роботи української редакції Texty.org.ua, яка інтегрує інструменти штучного інтелекту у процес створення власного контенту. Досліджено, як редакція визначає методологію для виявлення маніпулятивних прийомів, розрізняє гіпотези та факти, та перевіряє результати роботи власних натренованих алгоритмів.

Отримані результати свідчать, що поєднання інструментів ШІ з професійною експертизою журналістів забезпечує масштабованість аналізу без втрати змістовної точності, полегшує виявлення наративів та апеляції до конкретних емоцій, сприяє своєчасному виявленню організованих інформаційних кампаній.

Водночас метод не позбавлений ризиків, серед яких: упередження, хибні узагальнення, хибнопозитивні результати, необхідність самостійного тренування моделей та недосконалості самих алгоритмів у обробці природної мови і загальнолюдського контексту.

Практичний внесок полягає у сформульованих принципах редакційної інтеграції ШІ (прозорість, підзвітність, розділення ролей «модель–редактор»), а також у пропозиції процедурної моделі, придатної для впровадження у новинних організаціях із різними ресурсними можливостями. Теоретичний внесок виявляється у конкретизації умов, за яких ШІ виконує роль когнітивного підсилювача журналістської праці, зберігаючи стандарти професії. На прикладі української редакції показано, як ці умови реалізуються на практиці та які наслідки мають для якості аналітики й довіри аудиторії.

Ключові слова: *штучний інтелект, журналістика даних, журналістські практики, дезінформація, великі мовні моделі.*

Постановка проблеми. Стрімка інтеграція алгоритмів штучного інтелекту (ШІ) у журналістську практику поки що рідко переходить від рутинної автоматизації до повноцінного дослідницького застосування мовних моделей. Достатньо незначна кількість українських медіа використовує ШІ для досліджень та інтерпретації даних. На прикладах Texty.org.ua доведено, що ШІ у редакційних процесах може виконувати роль інструменту автоматизованої обробки текстів та виявлення і картографування маніпулятивних наративів. Водночас метод не позбавлений ризиків, серед яких: упередження, хибні узагальнення, хибнопозитивні результати, необхідність самостійного тренування моделей та недосконалості самих алгоритмів у обробці природної мови і загальнолюдського контексту.

Науково-практична значущість проблеми полягає в необхідності обґрунтувати й апробувати прозорий, масштабований і підзвітний алгоритм «модель + редактор», який поєднує збір і нормалізацію даних, анотацію, налаштування великої мовної моделі та людську верифікацію. Завданням є досягнення балансу між потенційними перева-

гами (масштабованість, точність, відтворюваність, економія часу) та ризиками (упередження, хибні узагальнення, хибнопозитивні спрацьовування, потреба додаткового тренування).

Аналіз останніх досліджень і публікацій. Недавні дослідження й публікації засвідчують активну інтеграцію штучного інтелекту у медіа та зміну підходів до створення і поширення новин. У вітчизняному дискурсі особливу увагу приділено адаптації інструментів ШІ до українського медіапростору та етичним наслідкам автоматизації; серед дослідників варто відзначити Ярошенко О. І. О. Є. Васьківську, О. Ситника, Т. М. Сидоренко, С. В. Машковець, К. Гнідковську, Д. Є. Мартиненка, які окреслюють ризики та можливості впровадження ШІ й вплив автоматизації на редакційні процеси.

На емпіричному рівні кейс Texty.org.ua демонструє застосування спеціально навченої мовної моделі для вимірювання маніпулятивності контенту українських Telegram-каналів, що показує практичне втілення людино-машинного підходу застосування ШІ в медіа. У закордонних студіях фокус зосереджено на глобальних трендах інтеграції ШІ у медіа, взаємодії журналістів з автоматизованими системами та оцінюванні ефективності моделей. Серед авторів – Льюїс, С. К., Вестлунд, О. Лю, Х., Ньюман, Н. Фернандеш, Е., Моро, С., та Кортез, П. Додатково загальносвітові тренди у новинних практиках і споживанні медіа надають щорічні огляди Reuters Institute [26], що підтверджує системний характер змін.

Постановка завдання. Метою статті є обґрунтувати можливості застосування ШІ як інструменту викриття (виявлення та інтерпретації) наративів, за допомогою аналізу лексичних та логічних маніпуляцій та автоматизованої обробки текстового контенту. А також – описати умови становлення такої практики та проілюструвати їх на матеріалах проєктів Texty.org.ua.

Виклад основного матеріалу. Розповсюдження інформаційних продуктів, створених за допомогою технології штучного інтелекту (ШІ) набуло значного масштабу з початку 2020-их років. Стрімкий прогрес великих мовних моделей (LLM) та широке впровадження генеративних інструментів ШІ (ChatGPT, Llama, Gemini, Copilot, Midjourney тощо) істотно трансформували журналістську практику. У цьому контексті поширення набув термін «журналістика, керована ШІ» (AI-driven journalism) [8].

Згідно зі звітом 2025 Digital News Report [26, ст. 11] значна частина медіа прагнуть використовувати штучний інтелект для кращої персоналізації новинного контенту. Також частина видавців бачать такі способи покращити досвід сприйняття новин за допомогою ШІ як підсумовування (27%), переклад історій різними мовами (24%), кращі рекомендації (21%) та використання чат-ботів для запитань про новини (18%).

Датою масового використання ШІ в медіа заведено вважати 2022 рік із появою нових інтерфейсів ШІ, так званих чат-ботів, які надавали доступ до моделей нейромереж, здатних продукувати тексти, зображення та відео на основі навчання на великих масивах даних [20]. Водночас, ще задовго до цієї дати відомо про значну кількість продуктів розроблених переважно великими західними медіакомпаніями, які намагались адаптувати ці інструменти для власних потреб редакцій. За глобальним опитуванням LSE JournalismAI, ще у 2019 р. половина опитаних новинних організацій використовувала ШІ [13, ст. 6]. У США Associated Press автоматизувало звіти про прибутки (згодом – і спортивні прев'ю), що вивільнило час журналістів [9, 10, 11]. Bloomberg застосовував інструмент Cyborg для швидких матеріалів з корпоративної звітності [17].

В Україні найновішим дослідженням як редакції використовують ШІ є опитування Інституту масової інформації (IMI) у 2024 році. За його результатами, 22% редакцій застосовують ШІ на постійній основі; ще 30% користуються ним зрідка; 20% узагалі не використовують; 16% припинили попереднє використання. Серед причин стриманості домінують переконання у неперевершеності «людської індивідуальності» (28%), брак знань і навичок роботи з ШІ (25%), побоювання хибної інформації (19%), а також сумніви щодо якості та ризик плагіату (9%) [3, 4].

Серед тих працівників медіа, що використовують ШІ, найчастіше використовуються такі функції:

- 62% – для розшифровки інтерв'ю;
- 31% – для створення індивідуального візуального контенту для матеріалів;
- 21% – виправлення помилок у текстах;
- 21% – написання статей, репортажів, пресрелізів та інших текстових матеріалів;
- 19% – генерування заголовків та лідів до текстів;
- 17% – для оптимізації журналістських текстів для різних аудиторій та платформ;
- 14% – для обробки та аналізу великих масивів даних;
- 12% – підготовка питань до інтерв'ю;

• 7% – збирання згадок в медіа та моніторингу репутації.

Як очевидно з результатів, переважна більшість медіа використовують ШІ для переважно рутинних операцій або ж у конкретному фрагменті підготовки журналістського продукту. Також частою практикою, не вказаною в опитуванні, є використання систем ШІ для озвучення текстів на сайтах [3]. У дослідженні IMI, 74% опитаних, які використовують інструменти штучного інтелекту, зазначили, що використовують ChatGPT, який є найпопулярнішим застосунком у цій сфері серед медійників.

Попри те, що українські медіа дедалі активніше інтегрують інструменти штучного інтелекту у повсякденні виробничі процедури – від транскрибування та перекладів до рутинного редагування і підготовки чернеток, – застосування ШІ для повноцінних дослідницьких проєктів, побудови й експлуатації власних мовних моделей або публікацій у форматі дата-журналістики на базі ШІ поки що має радше поодинокий характер. На тлі цієї тенденції Texty.org.ua вирізняються як редакція, що послідовно й системно використовує можливості великих мовних моделей у всьому циклі роботи. За словами Наталії Романишин, (NLP Researcher у Texty.org.ua), редакція активно використовує штучний інтелект протягом багатьох років, що є частиною їхньої редакційної практики, орієнтованої на дата-журналістику [6].

Особливої уваги заслуговує дослідження видання під назвою «Карусель емоцій», що стало одним з переможців міжнародного конкурсу з журналістики даних Sigma Awards 2025.

Дослідження вимірювало рівень маніпулятивності українських телеграм-каналів зо допомогою спеціально навченої мовної моделі [1]. Дизайн дослідження охоплював топ-100 каналів (50 у категорії «Новини і медіа» та 50 у категорії «Політика» за даними Telemetr.io станом на 1 вересня 2024 р.) (рис 1.).

Головним викликом для редакції було опрацювання коректної методологічної рамки дослідження. Як зазначає Наталії Романишин, стандарти та класична література з технік пропаганди та типові означення маніпуляцій, визначались ще у першій половині XX ст. та не повністю відповідають практикам поширеним у середовищі Telegram. Тому команда сформувала фокус-групу з журналістів, медіааналітиків, дослідників дезінформації та людей, які працюють з новинними стрічками в Telegram, а також користались напрацюваннями проєкту «Тактики пропаганди» від «Детектор медіа» [6].

У результаті було визначено 10 найбільш релевантних технік, які можна знайти в українському Telegram і які можна шукати майже без контексту. Техніки поділяються на емоційні (апеляції до страху, емоційна лексика, узагальнення) та маніпуляції аргументом (складніші, що вимагають контексту, «сам дурень!» (Whataboutism); вибіркова правда (Cherry picking) чи опудало (Straw man)).

Адаптація моделі відбувалась через ручну розмітку 9 512 публікацій командою з 14 анотаторів із різним профілем (журналісти, дата-аналітики, фахівці з комерційної анотації, маркетолог), що підвищило різноманіття інтерпретацій та якість інструкцій.

З подальшою редакторською рецензією провели донавчання моделі Meta-Llama-3.1-8B-Instruct на підмножині із 7 604 прикладів [1]. Корпус формувався протягом трьох літніх місяців 2024 року, що забезпечило достатній обсяг повідомлень від топ-100 каналів Telegram (за версію Telemetr). Для масштабування й здешевлення вартості обробки обрано підхід із моделлю відкритих ваг (open-weights), яку можна локально донавчати та багаторазово використовувати на власних комп'ютерах.

Формування корпусу даних вимагає різноманітної вибірки за часом, тематикою та джерелами. Редакція цілеспрямовано включала у вибірку як канали, відомі поширенням дезінформації, так і офіційні медіа, що дає змогу зіставляти наративи та риторичні практики в різних інформаційних екосистемах [6].

Вибір моделі з відкритими вагами зумовлено двома факторами: (а) потребою у спеціалізованій задачі (класифікація 10 технік), де загальномовні пропрієтарні LLM давали нижчу точність (~80% за пілотними тестами); (б) економічною доцільністю при обробці великих масивів повідомлень.

На етапі уточнення категоріального апарату фокус-група змінила початкове уявлення про «маніпуляцію» як суто негативне явище, додавши поняття «позитивної маніпуляції». Наприклад, повідомлення, спрямовані на підвищення морального стану аудиторії чи спроби емоційної залученості аудиторії до спортивних досягнень представників збірної тощо.

Донавчання моделі також дозволило внести певні корективи у інтерпретацію повідомлень великою мовною моделлю. Зокрема великі мовні моделі можуть хибно інтерпретувати повідо-

млення (наприклад, про повітряну тривогу як апеляцію до страху), тоді як для аудиторії це фактична інформація. Окремо обговорювалося виключення емоційно навантажених лексем (на кшталт «русня», «москаль») із переліку маніпулятивних індикаторів через їх контекстну інтерпретацію.

Підсумкова модель досягла загальної точності 86% і прецизійності 91% для класу «маніпулятивні публікації», класифікуючи десять технік, зокрема «навантажену мову», «блискучі узагальнення», «ейфорію», «апеляцію до страху» та «страх, непевність і сумніви». Для візуалізації результатів близькі за змістом техніки частково агрегували [1].

Отримані результати свідчать, що поширеність маніпуляцій суттєво варіює між каналами – від менш як 2% до понад 90% повідомлень [1, 22].

Серед деструктивних стратегій найбільш типовими виявилися апеляції до страху та навіювання сумнівів, що корелює з практиками проросійської пропаганди в українському інформаційному просторі [1] (рис. 2).

Кейс продемонстрував можливість автоматизованого аналізу маніпуляцій та оманливих наративів.

Водночас такий метод містить свої обмеження та недоліки. Важливо розуміти, що емоційна лексика не дорівнює порушенням журналістських стандартів чи свідомо шкідливій маніпуляції, однак дозволяє проаналізувати джерело інформації на предмет об'єктивності. (У застосованій таксономії емоційна мова розглядається як окрема маніпулятивна техніка й водночас не обов'язково оцінюється як нормативне порушення з огляду на жанр і контекст телеграм-дискурсу). І навпаки, іноді пропагандисти мімікрують під виважене ЗМІ й не використовують маніпулятивної лексики.

Також попри високу точність, модель не завжди могла розпізнавати сарказм, що підкреслює наразі неідеальну інтерпретацію тексту за допомогою ШІ.

Досвід апробований у проєкті «Карусель емоцій» (збір каналів за Telemetr, ручна розмітка, донавчання інструктивної моделі та валі-

Рис. 1. Рівень маніпулятивності за 10 показниками в різних груп каналів

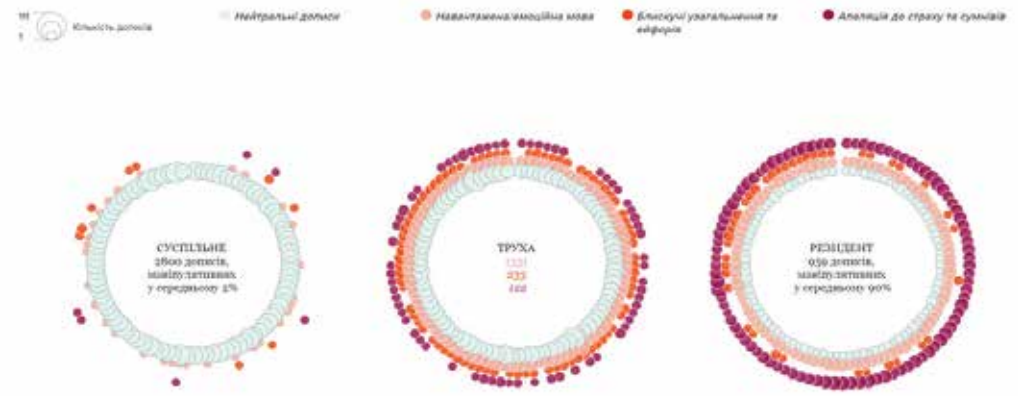

Рис. 2. Деякі приклади візуалізації емоційно маніпулятивної лексики у телеграм-каналів

дація) продемонстрував можливість повторного використання й адаптації вже напрацьованих процедур.

Для невеликих обсягів даних, менш складних задач (проста класифікація, пошук інформації) або коли актуальність проблеми недовговічна редакція використовує запити до пропрієтарних моделей OpenAI. Зокрема для аналізу скоординованої кампанії дискредитації НАБУ/САП улітку 2025 р.

Вибірка 100 найпопулярніших каналів категорії «Новини та медіа» за даними Telemetr.io станом на 23 липня 2025 р., пройшла подальшу фільтрацію до 58 каналів і тематичний збір повідомлень за із ключовими словами про НАБУ/САП [5]. Кожен допис автоматично класифікувався моделлю GPT-4.1 за трьома категоріями (oppositional/supportive/neutral); для подальшого аналізу обрано 24 канали, які принаймні двічі поширювали «oppositional»-повідомлення, – загалом 246 дописів у вибірці (рис. 3).

У часовому профілі кампанія синхронізувалася з правоохоронними діями щодо працівників НАБУ (кульмінація – 21 липня) та з проходженням законопроєкту № 12414 і його оперативним підписанням 22 липня, що свідчить про узгодженість інформаційних та політико-правових подій [25].

На рівні наративів і джерел легітимації ключову роль відігравали коментарі політконсультантів Олега Постернака (14 цитувань) і Михайла Шнайдера (8), які приписували НАБУ «філіацію» з «силовим органом країни-агресора» або неспроможність громадського сектору розпізнавати «держзраду»; ці меседжі підсилювалися повторюваними формулами («розпустити НАБУ», «НАБУ проїдає гроші» тощо).

У підсумку це дозволяє побачити узгоджену поведінку та здатність масштабувати недовірливі фрейми щодо інституції, тоді як застосований редакцією підхід відтворює і розширює попередні алгоритмічні напрацювання з «Каруселі емоцій».

Подібний аналіз (з використанням власної навченої моделі) використовувався і в інших кейсах, зокрема маніпулятивних наративів зі зриву мобілізації запущених у соцмережах авторкою під псевдонімом Соломія Українець [2].

Подібні дослідження робить також британська фактчекерська організація Full Fact. Традиційні моделі машинного навчання добре працюють для класифікації контенту за заздалегідь визначеними рубриками. Один із компонентів Full Fact – тематичний класифікатор, що визначає одну або кілька тем кожної новинної публікації. Команда сформувала фіксований перелік із 17 тематик (зокрема охорона здоров'я, політика, освіта) та навчила модель на 10 тис. розмічених статей. Після цього класифікатор стабільно й передбачувано ставить мітки новим матеріалам – це полегшує фільтрацію та відбір [19]. В організації наголошують, що генеративні моделі навчаються на мільярдах речень без фіксованих міток, тож мають сильну «мовну інтуїцію», але не гарантують глибинного розуміння істинності. Вони корисні для запитань на кшталт «Хто і що тут стверджує?» або «Чи рівнозначні ці два твердження?», але не відповідають на питання «Чи це

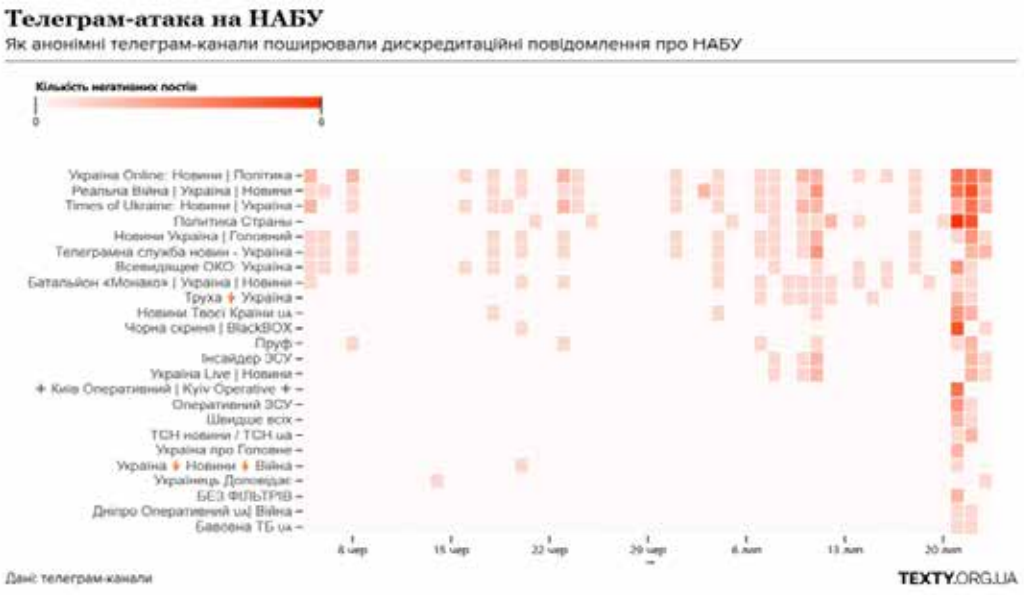

Рис. 3. Візуалізація наративної кампанії щодо дискредитації НАБУ

правда?», адже істинність залежить від зовнішніх знань і перевірки фактів.

Іншим вартим уваги кейсом є інструмент Haystacker, який впровадила редакція The Washington Post. У серпні 2024 видання вперше опублікувало статтю, побудовану на роботі нового інструменту штучного інтелекту. У показовій публікації редакція проаналізувала 700+ політичних роликів про імміграцію і виявила, що близько п'ятої частини використовують застарілі або вирвані з контексту кадри, спонсоровані Республіканською партією. Haystacker дозволяє журналістам переглядати великі набори даних – відео, фото чи текст – щоб знаходити варті уваги тенденції чи закономірності [18]. Інструмент, на створення якого знадобилося понад рік, використовується переважно командами візуальної криміналістики та журналістики даних The Washington Post. Хоча Haystacker має на меті допомогти журналістам вивчати великі набори даних, його також можна використовувати для швидкого узагальнення відеоматеріалів.

Однак, як було згадано раніше, використання власних інструментів на базі ШІ є відносно нетиповою практикою для медіа. Серед причин чому інші медіа не створюють власні інструменти, Наталія Романишин виділяє декілька основних:

Редакційна стратегія. Не для всіх медіа це пріоритет; Texty.org.ua спеціалізуються на дата-журналістиці, де користь від власних рішень максимальна.

Відсутність технічних навичок: Більшість працівників медіа не мають необхідних технічних навичок для розробки власних інструментів.

Час і витрати. Розроблення повного циклу (корпус → розмітка → навчання → інтеграція) є трудомістким; Для дослідження «Карусель емоцій» цикл становив ≈4 місяці.

Альтернативи співпраці. Частина редакцій обирає партнерства з аналітичними компаніями або дослідницьким організаціям, делегуючи їм збір і обробку великих масивів даних.

Висновки. Інтеграція ШІ в редакційні процеси трансформуватиме роботу медіа та ролі журналістів: алгоритми здатні автоматизувати рутинні операції (розшифровки, базове резюмування, шаблонні тексти), звільняючи час для аналітики та репортерської роботи, водночас посилюючи ризики упереджень, помилок, втрати прозорості та порушення авторських прав [25, 26]. Вважається, що розробки в галузі ШІ дедалі більше впливатимуть на медіаекосистеми в цілому.

Поза іншими аспектами, помітний значний вплив ШІ на журналістику даних (data journalism). Інструменти машинного навчання та NLP дедалі частіше стають частиною стандартного пайплайну дата-журналістики: від збору та очищення даних до класифікації, вилучення сутностей, підсумовування та візуалізації.

Досвід Texty.org.ua демонструє потенціал ШІ як інструменту для аналізу великих обсягів текстового контенту та виявлення наративів. Однак, важливо усвідомлювати обмеження ШІ та необхідність ретельної перевірки результатів за допомогою людини. А також усвідомлювати, що мовні моделі мають власні упередження та системні викривлення [15].

У всіх проєктах підкреслюється важливість перевірки результатів, отриманих за допомогою ШІ. Наприклад, у в проєкті «Карусель Емоцій» використовувалися ручна двоетапна розмітка даних та їх перевірка, тестування моделі на обмежених обсягах та залучення додаткових експертів. Наразі людський фактор залишається критичним для забезпечення точності та надійності результатів.Тому поєднання можливостей ШІ з експертизою журналістів є ключовим для ефективного використання потенціалу великих мовних моделей у редакційній практиці.

Поступово ШІ у журналістиці може вийти за межі помічника у рутинних завданнях, а стати інструментом для проведення власних досліджень, пошуку аномалій у статистичних даних, автоматичного аналізу маніпулятивних наративів тощо.

Результати автоматизованого аналізу потребують обов'язкової людської верифікації, оскільки ймовірна втрата контексту (наприклад, смайл емодзі, який сміється у репості може сигналізувати сарказм, а не згоду з позицією). Тому остаточне рішення щодо верифікації та інтерпретації даних у Texty.org завжди ухвалює людина.

Список літератури:

1. Карусель емоцій. Рівень маніпулятивності мови українських телеграм-каналів / Н. Кельм та ін. Texty.org.ua – статті та журналістика даних для людей – Тексти.org.ua. URL: https://texty.org.ua/projects/113607/karusel-emocij-riven-manipulyatyvnosti-movy-ukrayinskyh-telehram-kanaliv/ (дата звернення: 23.10.2025).

2. Клікбейт на крові. Що не так із віршами Соломії Українець і як їх використовують китайці та росіяни / І. Гадзинська та ін. Texty.org.ua URL:https://texty.org.ua/articles/114381/klikbejt-na-krovi-sho-ne-tak-iz-virshamy-solomiyi-ukrayinec-i-yak-yih-vykorystovuyut-kytajci-ta-rosiyany/ (дата звернення: 23.10.2025).

3. Машкова Я. Опитування IMI: 22% українських редакцій використовують штучний інтелект на постійній основі. Інститут масової інформації. URL: https://imi.org.ua/news/opytuvannya-imi-22-ukrayinskyh-redaktsij-vykorystovuyut-shtuchnyj-intelekt-na-postijnij-osnovi-i62243 (дата звернення: 23.10.2025).

4. Машкова Я. Українські медіа та штучний інтелект. Як редакції залучають ШІ для створення контенту?. Інститут масової інформації. URL: https://imi.org.ua/monitorings/ukrayinski-media-ta-shtuchnyj-intelekt-yak-redaktsiyi-zaluchayut-shi-dlya-stvorennya-kontentu-i62217 (дата звернення: 23.10.2025).dat

5. «Розігнати НАБУ». Як пул анонімних телеграм-каналів злагоджено підтримав ініціативу президента / І. Гадзинська та ін. Texty.org.ua. URL:https://texty.org.ua/articles/115560/rozihnaty-nabu-yak-pul-anonimnyh-telehram-kanaliv-zlahodzheno-pidtrymav-iniciatyvu-prezydenta (дата звернення: 10.10.2025).

6. Романишин Наталія (NLP Researcher, Texty.org.ua). Інтерв'ю досліднику / інтерв'юер: Тартачний О. В.; розшифровка: Тартачний О. В. [Неопубліковане інтерв'ю, розшифровка]. Київ, 2025. 34 с.

7. Що думає ШІ про Україну? Дослідження упереджень великих мовних моделей / Є. Костюк та ін. Texty.org.ua – статті та журналістика даних для людей – Тексти.org.ua. URL: https://texty.org.ua/projects/115548/sho-dumaye-shi-pro-ukrayinu-doslidzhennya-uperedzhen-velykyh-movnyh-modelej/ (дата звернення: 23.10.2025).

8. A data-centric approach for ethical and trustworthy AI in journalism / L. Dierickx et al. *Ethics and information technology*. 2024. Vol. 26, no. 4. URL: https://doi.org/10.1007/s10676-024-09801-6 (date of access: 23.10.2025).

9. Associated Press. A leap forward in quarterly earnings stories. Associated Press. URL: https://www.ap.org/the-definitive-source/announcements/a-leap-forward-in-quarterly-earnings-stories/ (date of access: 10.10.2025).

10. AP expands Minor League Baseball coverage | The Associated Press. The Associated Press. URL: http://www.ap.org/media-center/press-releases/2016/ap-expands-minor-league-baseball-coverage/ (date of access: 10.10.2025).

11. Associated Press. Updates to generative AI standards. Associated Press. URL: https://www.ap.org/the-definitive-source/behind-the-news/updates-to-generative-ai-standards/ (date of access: 10.10.2025).

12. Standards around generative AI | The Associated Press. The Associated Press. URL: https://www.ap.org/the-definitive-source/behind-the-news/standards-around-generative-ai/ (date of access: 23.10.2025).

13. Beckett C. New powers, new responsibilities A global survey of journalism and artificial intelligence. London : The London School of Economics and Political Science, 2019. 111 p. URL: https://eprints.lse.ac.uk/129641/1/New_powers_new_responsibilities_The_Journalism_AI_report_English.pdf (date of access: 10.10.2025).

14. Carlson M. The robotic reporter. *Digital journalism*. 2014. Vol. 3, no. 3. P. 416–431. URL: https://doi.org/10.1080/21670811.2014.976412 (date of access: 23.10.2025).

15. Entman R. M. Framing: toward clarification of a fractured paradigm. *Journal of communication*. 1993. Vol. 43, no. 4. P. 51–58. URL: https://doi.org/10.1111/j.1460-2466.1993.tb01304.x (date of access: 23.10.2025).

16. European Journalism Centre. The state of data journalism survey 2023. Maastricht : European Journalism Centre (EJC), 2023. 44 p. URL: https://aprensamalaga.com/wp-content/uploads/2024/04/The-State-of-Data-Journalism-2023-pdf-report.pdf (date of access: 10.10.2025).

17. Fischer S. Rise of the content bots. Axios. URL: https://www.axios.com/2019/02/07/rise-of-the-content-bots (date of access: 10.10.2025).

18. Fischer S. New Washington Post AI tool sifts massive data sets. Axios. URL: https://www.axios.com/2024/08/20/washington-post-ai-tool-data (date of access: 10.10.2025).

19. How AI can help fact checkers – Full Fact. Full Fact. URL: https://fullfact.org/blog/2025/feb/how-ai-can-help-fact-checkers/ (date of access: 23.10.2025).

20. Introducing ChatGPT. openai.com/. URL: https://openai.com/index/chatgpt (date of access: 10.10.2025).

21. Lewis S. C., Westlund O. Actors, Actants, Audiences, and Activities in Cross-Media News Work. Digital Journalism. 2014. Vol. 3, no. 1. P. 19–37. URL: https://doi.org/10.1080/21670811.2014.927986 (date of access: 23.10.2025).

22. Global Investigative Journalism Network. From flammable buildings to slavery's hidden legacy to tainted groundwater: projects from 10 countries win gijn's 2025 sigma awards. Global Investigative Journalism Network. URL: https://gijn.org/stories/gijn-winners-2025-sigma-awards/ (date of access: 23.10.2025).

23. Hebura et al. Carousel of emotions. manipulation level of ukrainian telegram channels / Texty.org.ua. URL: https://texty.org.ua/projects/113693/carousel-of-emotions-manipulation-level-of-ukrainian-telegram-channels/ (date of access: 10.10.2025).

24. McCombs M. E., Shaw D. L. The agenda-setting function of mass media1 2. *The agenda setting journal*. 2017. Vol. 1, no. 2. P. 105–116. URL: https://doi.org/10.1075/asj.1.2.02mcc (date of access: 23.10.2025).

25. Newman N. Journalism, media, and technology trends and predictions 2024. Reuters Institute for the Study of Journalism. URL: https://reutersinstitute.politics.ox.ac.uk/journalism-media-and-technology-trends-and-predictions-2024 (date of access: 23.10.2025).

26. Newman N.; Fletcher R.; Robertson C. T.; et al. Digital News Report 2025. Reuters Institute for the Study of Journalism. URL: https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2025 (date of access: 10.10.2025)

27. The Washington Post. Republicans flood TV with misleading ads about immigration, border. The Washington Post. URL: https://www.washingtonpost.com/elections/interactive/2024/republican-campaign-ads-immigration-border-security/ (date of access: 10.10.2025).

Tartachnyi O. V. DETECTING NARRATIVES AND AUTOMATED TEXT PROCESSING WITH AI: APPLICATIONS AND THE EVOLUTION OF NEWSROOM PRACTICE (A CASE STUDY OF TEXTY.ORG.UA)

The article offers a practical reflection on the use of large language models (LLMs) in journalism for the automated processing of texts and for detecting and mapping manipulative narratives.

The core problem is that most newsrooms integrate AI at the level of transcription, translation, or technical copyediting, but seldom employ it as a tool for topic discovery or data interpretation. Our study focuses on the experience of a Ukrainian newsroom that systematically combines the capabilities of large language models with editorial expertise. Using cases from Texty.org.ua, we show that large language models can recognize common manipulation techniques even in short posts on social platforms.

The article combines an analytical review of relevant approaches used by foreign newsrooms with an in-depth empirical examination of the operating principles of the Ukrainian newsroom Texty.org.ua, which integrates AI tools into the process of producing its own content. We explore how the newsroom defines a methodology for identifying manipulative techniques, distinguishes between hypotheses and facts, and validates the outputs of its own trained algorithms.

The findings indicate that combining AI tools with journalists' professional expertise enables scalable analysis without sacrificing substantive accuracy, facilitates the detection of narratives and appeals to specific emotions, and supports the timely identification of organized information campaigns. At the same time, the method is not free of risks, including bias, faulty generalizations, false positives, the need to train models in-house, and the imperfections of the algorithms themselves in processing natural language and broader human context. The practical contribution lies in the articulated principles for editorial AI integration (transparency, accountability, and clear separation of the "model–editor" roles), as well as in proposing a procedural model suitable for implementation in news organizations with varying resource capacities. The theoretical contribution specifies the conditions under which AI acts as a cognitive amplifier of journalistic work while upholding professional standards. Using the Ukrainian newsroom as an example, we demonstrate how these conditions are realized in practice and what implications they have for the quality of analysis and audience trust.

Key words: *artificial intelligence, data journalism, journalistic practices, disinformation, large language models.*

Дата надходження статті: 11.10.2025
Дата прийняття статті: 10.11.2025
Опубліковано: 29.12.2025