УДК 070:004
DOI https://doi.org/10.32782/2710-4656/2025.5.2/60

Каніщев С. О.
Київський національний університет імені Тараса Шевченка

ПРОБЛЕМИ ГЕНЕРАЦІЇ КОНТЕНТУ ШІ У ЖУРНАЛІСТИЦІ: ЕПІСТЕМІЧНІ, ЕТИЧНІ ТА ОРГАНІЗАЦІЙНІ ВИМІРИ

У статті запропоновано комплексну Систему управління використанням штучного інтелекту в редакціях (СУШІР: нову, процесно-орієнтовану модель, розроблену для мінімізації епістемічних, етичних та організаційних ризиків, пов'язаних із впровадженням генеративного штучного інтелекту (ШІ) в журналістську практику. Актуальність дослідження зумовлена стрімкою, але часто несистематичною інтеграцією генеративних моделей у редакційні процеси, що створює загрози для достовірності, прозорості та підзвітності медіа. Метою статті є розробка практичного інструментарію, що дозволяє редакціям структурувати роботу з ШІ, зберігаючи при цьому ключові журналістські стандарти. Методологія дослідження базується на якісному аналізі та синтезі наявних редакційних політик провідних світових медіа, огляді актуальних наукових публікацій щодо ризиків ШІ та застосуванні принципів процесного управління й ризик-менеджменту до журналістського виробничого циклу. Результати роботи представлені у вигляді СУШІР, що включає кілька ключових компонентів: 1) детальний операційний протокол: «запит, генерація, верифікація, маркування, публікація, постконтроль». 2) матрицю оцінки ризиків для різних журналістських жанрів. 3) рекомендації щодо впровадження технологічних рішень, таких як генерація тексту на основі заздалегідь перевірених джерел для зменшення фактологічних помилок та використання стандартів для верифікації походження медіаконтенту. 4) набір редакційних політик щодо маркування, фактчекінгу та визначення «червоних ліній» для використання ШІ. Особливу увагу приділено застосуванню системи в українському контексті, де умови інформаційної війни висувають підвищені вимоги до інформаційної безпеки та протидії дезінформації. Наукова новизна полягає у переході від фрагментарних етичних рекомендацій до цілісної, операційної системи управління, що поєднує технологічні, організаційні та етичні запобіжники. Практична цінність роботи полягає у формуванні готової до впровадження моделі, яка стандартизує роботу з генеративним ШІ, підвищує керованість ризиків та сприяє збереженню довіри аудиторії.

Ключові слова: *генеративний ШІ, журналістика, етика, фактчекінг, дезінформація, редакційні політики, СУШІР.*

Постановка проблеми. Стрімка інтеграція генеративних моделей штучного інтелекту (ШІ) у редакційні процеси створює фундаментальний виклик для журналістики як соціального інституту. Цей процес відбувається на тлі вже наявних криз: падіння довіри до медіа, економічної нестабільності та поляризації інформаційного простору. Проблема полягає не стільки в самій технології, скільки у відсутності системного підходу до управління ризиками, що її супроводжують. Використання ШІ без належних операційних рамок, часто на рівні індивідуальних експериментів журналістів, призводить до епістемічних, етичних та організаційних інцидентів. Такі інциденти, навіть поодинокі, здатні підривати достовірність контенту та руйнувати довіру аудиторії, відновлення якої є надзвичайно складним завданням. Таким чином, виникає гостра необхідність у розробці комплексної системи, яка б дозволила медіа використовувати переваги ШІ, не жертвуючи при цьому професійними стандартами та своєю суспільною відповідальністю.

Аналіз останніх досліджень і публікацій. Науковий дискурс навколо ШІ в журналістиці активно розвивається, фокусуючись переважно на ідентифікації та класифікації ризиків. Дослідники виокремлюють епістемічні загрози, зокрема системну природу «галюцинацій» (фактологічних помилок) та вплив когнітивних упереджень, таких як автоматизаційна упередженість [5]. Значна увага приділяється етичним дилемам, пов'язаним із прозорістю, підзвітністю та упередженістю алгоритмів [1]. Окремим напрямом є аналіз правової невизначеності, насамперед у сфері авторського права щодо навчальних даних та згенерованого контенту [14; 20]. Численні медійні

організації та професійні об'єднання пропонують етичні кодекси та рекомендації, що наголошують на необхідності людського контролю та маркування контенту [6; 7]. В українському контексті державні органи також розробили настанови, акцентуючи на підвищених ризиках в умовах інформаційної війни [2; 3]. Попри значний масив напрацювань, існує розрив між теоретичним осмисленням проблем та наявністю практичних, інтегрованих моделей управління. які б поєднували процесні, технологічні та політичні аспекти в єдину систему для щоденної роботи редакції.

Постановка завдання. Мета статті: розробити та обґрунтувати комплексну Систему управління використанням ШІ в редакціях (СУШІР) як процесно-орієнтовану модель, що систематизує впровадження генеративних технологій та мінімізує пов'язані з ними ризики, забезпечуючи дотримання журналістських стандартів.

Виклад основного матеріалу. Метрики та контрольні точки. Ключові показники впровадження СУШІР: точність (частка підтверджених тверджень), повнота (покриття обов'язкових фактів), час підготовки матеріалу, частота і тяжкість інцидентів (від стилістичних хиб до правових претензій), частка матеріалів із коректним маркуванням участі ШІ, частка матеріалів із подвійною перевіркою фактів у чутливих темах. Контрольні точки: передпублікаційний контрольний список; подвійна перевірка фактів для чутливих тем; попередній юридичний перегляд для високоризикових матеріалів; журналювання джерел кожного факту; післяпублікаційний моніторинг і журнал змін; щоквартальний аудит політик.

1. Епістемічні ризики: знання, яке імітує себе саме. Фундаментальний епістемічний ризик генеративного ШІ полягає в його природі: модель не «знає» фактів у людському розумінні, а передбачає найбільш статистично ймовірні текстові послідовності на основі навчальних даних. Її синтаксична впевненість легко вводить в оману, створюючи ілюзію істинності. Системи на базі великих мовних моделей продукують правдоподібні форми, але не гарантують істинність змісту. Так звані «галюцинації», тобто іншими словами фабрикація фактів, джерел, цитат є не випадковою похибкою, а системною властивістю цього процесу [6]. Для новинних жанрів, де точність є абсолютною вимогою, це становить пряму загрозу довірі аудиторії. Цю проблему посилює автоматизаційна упередженість (automation bias): «гладкий», ритмічно збудований і стилістично переконливий текст психологічно сприймається як більш достовірний [5]. В умовах жорстких дедлайнів редакції ризикують приймати машинні версії фактів без належної перевірки, покладаючись на їхню зовнішню переконливість. Таким чином, будь-який сирий вихід моделі має розглядатися виключно як чернетка, що потребує незалежної людської верифікації за першоджерелами.

2. Лінгвістично-дискурсивні ризики: усереднення стилю та втрата авторського голосу. Хоча великі мовні моделі чудово відтворюють «середню норму» мови, вони слабкі в передачі оригінального авторського голосу та стилю. Надмірне використання генеративних інструментів призводить до стилістичної конвергенції: тон матеріалів вирівнюється, мовні маркери стираються, а дискурсивна різноманітність медіапростору зменшується. У першу чергу це представляє загрозу для найцінішого нематеріального активу видання, його унікального редакційного голосу. Розмивання цієї ідентичності через усереднений машинний стиль є не просто естетичною проблемою, а прямою загрозою бізнес-моделі якісної журналістики [7]. Іншим ризиком є псевдозв'язність, тобто здатність моделей генерувати плавні, граматично коректні, але концептуально порожні тексти. Аргументи можуть повторюватися, причинно-наслідкові зв'язки підмінятися асоціативними, а часові площини змішуватися без логічних переходів. Це створює

Таблиця 1

Матриця ризиків СУШІР

Категорія ризику	Опис	Міри контролю
Епістемічний	Галюцинації, помилкові узагальнення.	Подвійна перевірка фактів; посилання на джерела; заборона неперевірених тверджень.
Етичний	Приховане використання ШІ, маніпуляції зображеннями.	Маркування участі ШІ; етичний перегляд; перевірка різноманітності та недискримінаційності.
Правовий	Авторське право, персональні дані, дифамація.	Попередній юридичний перегляд; білі списки; ліцензування; аудит журналів дій.
Організаційний	Розмитість відповідальності, залежність від інструмента.	Матриця ролей і відповідальності; політика журналювання; резервні процеси; навчання персоналу.

ілюзію змістовності, маскуючи відсутність глибокого аналізу. Для українського медіаполя особливу небезпеку становить автоматичний переклад, що несе ризик прихованих смислових зсувів та непомітного калькування англомовних синтаксичних структур.

3. Етичні ризики: межі авторства і відповідальності. Інтеграція моделей ШІ робить видимими етичні межі журналістики, такі як авторство, прозорість і відсутність дискримінації. Коли машина допомагає створювати текст чи візуальні матеріали, ланцюг відповідальності легко розмивається. Тому підзвітність має бути чітко прив'язана до людини, що ухвалює редакційні рішення, а ступінь участі ШІ відкрито пояснений читачеві. Суть не в заборонах, а у прозорих процедурах: хто сформулював завдання, що саме автоматизовано, що перевірено людиною, які джерела використано. Маркування участі ШІ таким чином можна розглядати як елемент чесності, навіть якщо воно інколи сприймається як сигнал меншої надійності. Це упередження долається не відмовою від маркування, а зрозумілим формулюванням: «інструменти ШІ застосовано для узагальнення матеріалів; усі факти перевірено редактором за джерелами». Таке пояснення демонструє процес людського контролю й підтримує довіру. Окремою площиною є захист даних і недискримінація. Запити до моделей можуть містити персональну чи службову інформацію, отже потрібні корпоративні налаштування, чіткі обмеження і навчання персоналу інформаційній гігієні. Щоб уникати вбудованих у моделі упереджень, варто запроваджувати етичний перегляд чутливих матеріалів і перевірку аспектів представлення тем і героїв. Це робить використання ШІ сумісним із цінностями професії.

4. Правові ризики: авторське право, дифамація, регулювання. Правові ризики пов'язані насамперед із авторським правом, відповідальністю за зміст і змінами у регулюванні. Навчання моделей на захищених творах і використання згенерованих фрагментів створюють «сіру зону» похідних творів і складні питання ліцензування, але правова охорона можлива лише за наявності істотного людського внеску. Для редакцій це означає документувати творчі дії людини, уникати відтворення чужих стилів і, за можливості, спиратися на ліцензовані чи власні корпуси даних у чутливих кейсах. Юридична відповідальність за публікацію залишається за видавцем незалежно від того, як сталася помилка. Тому багаторівнева перевірка фактів і попередній юридичний перегляд для матеріалів підвищеного ризику має бути стандартом роботи редакції. Паралельно варто стежити за міжнародними підходами до регулювання ШІ: ризик-орієнтована модель ЄС, більш фрагментована американська практика та рамкові документи Ради Європи задають вимоги до прозорості й підзвітності, які впливатимуть і на українські медіа.

5. Операційний протокол для редакції (СУШІР): від запиту до пост контролю. Для перетворення ризиків на керовані процеси необхідно впровадити формалізований операційний протокол. Пропонується модель СУШІР (Система Управління використанням ШІ в Редакціях), що охоплює весь життєвий цикл створення контенту за допомогою генеративного ШІ. Цей протокол складається з дев'яти послідовних етапів, які утворюють замкнений цикл постійного вдосконалення:

- Запит: Журналіст чітко формулює завдання: що саме потрібно згенерувати, з якою метою, і які обмеження слід застосувати.
- Підготовка корпусу (за потреби): Для завдань, що вимагають високої точності, використовується підхід генерації на основі локального, перевіреного та ліцензованого корпусу джерел.
- Генерація: Процес генерації документується: фіксуються версія моделі, ключові параметри та повний текст запиту (промпту).
- Верифікація: Ключовий етап людського контролю. Редактор перевіряє всі факти, дати, власні назви, цитати та логічну когерентність тексту.
- Юридичний/етичний огляд: Матеріали на чутливі теми або з високим правовим ризиком передаються на розгляд юристу або етичному редактору.
- Маркування: До матеріалу додається чітке та зрозуміле повідомлення про роль ШІ у його створенні.
- Публікація: Матеріал публікується за підписом відповідальної людини (автора/редактора).
- Постконтроль: Після публікації здійснюється моніторинг реакції аудиторії та оперативне внесення виправлень.
- Аудит: Регулярний перегляд кейсів використання ШІ, аналіз помилок та оновлення внутрішніх політик.

Цей процес візуалізовано на схемі нижче.

6. Застосування у середовищі високих ризиків: український контекст. Український медіапростір функціонує в умовах повномасштабної війни, що робить його середовищем із надзви-

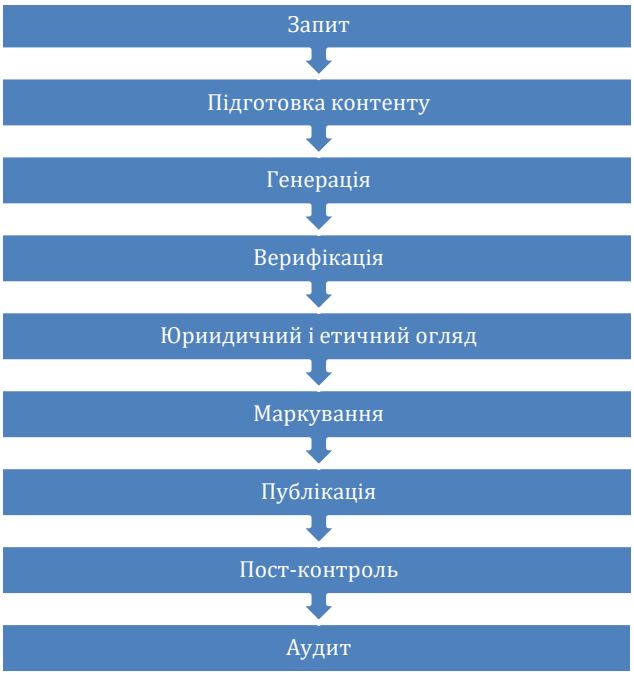

Рис. 1. Модель робочого циклу «СУШІР» (Система Управління використанням ШІ в Редакціях)

чайно високими ризиками. Російська Федерація активно використовує ШІ для автоматизації та масштабування дезінформаційних кампаній [12; 13]. ШІ-інструменти застосовуються для створення фейкових новин, маніпулятивних наративів та координації мереж ботів у соціальних мережах. У цьому контексті впровадження СУШІР набуває не лише етичного, а й стратегічного значення для інформаційної безпеки. Запропонована система слугує захисним механізмом, що підвищує стійкість редакцій до інформаційних атак. Пріоритет людської верифікації та використання технологій генерації на основі перевірених українських джерел допомагають запобігти несвідомому тиражуванню ворожих наративів. Суворі політики щодо безпеки даних унеможливлюють витік чутливої інформації. Таким чином, СУШІР стає інструментом не лише для забезпечення якості, а й для захисту інформаційного суверенітету, що узгоджується з рекомендаціями українських регуляторних органів [4].

Висновки. Генеративний ШІ є потужною технологією, що несе як значні можливості для оптимізації редакційних процесів, так і системні ризики для фундаментальних засад журналістики. Проблеми, пов'язані з його використанням, мають багатовимірний характер, охоплюючи епістемічні, етичні, правові та організаційні аспекти.

Наукова новизна цієї роботи полягає у пропозиції цілісної, процесно-орієнтованої Системи управління використанням ШІ в редакціях (СУШІР). На відміну від існуючих фрагментарних рекомендацій, СУШІР є комплексною операційною моделлю, що поєднує стратегічні принципи, покроковий протокол, технологічні запобіжники та набір конкретних редакційних політик. Ключовим результатом дослідження є обґрунтування того, що керованість ризиків ШІ досягається не через спроби усунути недоліки технології, а через побудову надійних організаційних процесів. Принцип «ШІ під контролем людини» перестає бути абстрактною декларацією і перетворюється на набір конкретних процедур верифікації, підзвітності та аудиту. Для українського медіапростору, що функціонує в умовах інформаційної війни, впровадження такої системи є не лише питанням професійної етики, а й елементом інформаційної безпеки. Перспективи подальших досліджень полягають в емпіричній апробації запропонованої системи в реальних редакціях, кількісній оцінці її впливу на якість контенту та рівень довіри аудиторії. Запропонована модель СУШІР поєднує продуктивність використання генеративного ШІ з епістемічними та етичними вимогами журналістики. Подальші дослідження варто спрямувати на кількісну оцінку ефектів (скорочення часу підготовки матеріалів, динаміка точності/повноти, зниження інцидентів), порівняння редакцій різного масштабу та інтеграцію автоматизованих інструментів перевірки мультимедійних даних.

Список літератури:

1. Авторське право на об’єкт, створений штучним інтелектом. ЛІГА:ЗАКОН, 2023. URL: https://jurliga.ligazakon.net/analitycs/225383_avtorsk-prava-na-obkt-stvoreniy-shtuchnim-ntelektom (дата звернення: 21.10.2025)

2. Рекомендації Комісії з журналістської етики щодо використання штучного інтелекту для створення журналістських матеріалів. Комісія з журналістської етики, 2023. URL: https://cje.org.ua/statements/rekomendatsii-komisii-z-zhurnalistskoi-etyky-shchodo-vykorystannia-shtuchnoho-intelektu-dlia-stvorennia-zhurnalistskykh-materialiv/ (дата звернення: 21.10.2025)

3. Як відповідально використовувати штучний інтелект: розробили рекомендації для медіа. Міністерство цифрової трансформації України, 2024. URL: https://thedigital.gov.ua/news/technologies/yak-vidpovidalno-vikoristovuvati-shtuchniy-intelekt-rozrobili-rekomendatsii-dlya-media (дата звернення: 21.10.2025)

4. AI-Driven Disinformation in the Russia-Ukraine War. Small Wars Journal, 2025. URL: https://smallwarsjournal.com/2025/10/02/ai-driven-disinformation/ (дата звернення: 21.10.2025)

5. Ali M. A., та ін. AI, Ethics, and Cognitive Bias: An LLM-Based Synthetic Simulation. *Journal of Media and Communication*, 2025, 1(1), 3. URL: 3390/journalmedia1010003. URL: .3390/journalmedia1010003 (дата звернення: 21.10.2025). DOI: 10.

6. Calvo-Rubio L. M., Ufarte-Ruiz M. J., Yezers'ka L. O. Ethical codes on the use of artificial intelligence in newsrooms: a comparative of the proposals of international media. *Profesional de la información*, 2024, 33(4), e330404. 2024.jul.04. 2024.jul.04 URL: 3145/epi.2024.jul.04. 2024.jul.04 URL: .3145/epi.2024.jul.04 (дата звернення: 21.10.2025). DOI: 10.

7. Diakopoulos N., та ін. Digital Newsroom Transformation: A Systematic Review of the Impact of Artificial Intelligence on Journalistic Practices, News Narratives, and Ethical Challenges. *Journal of Media and Communication*, 2024, 5(4), 97. URL: 3390/journalmedia5040097. (дата звернення: 21.10.2025).

8. Lewis S. C., Guzman A. L., Schmidt T. R. Generative AI and the disruptive challenge to journalism: an institutional analysis. Communication and the Public, 2025, 1(9), 1–20. URL: 1177/20570473251369305. (дата звернення: 21.10.2025).

9. Liu Q. Generative AI and Journalism Ethics: Controversies over ChatGPT. *Journal of Information, Technology and Policy*, 2025, 3(1), 1–6. 2025.346. 2025.346 URL: 62836/jitp.2025.346. 2025.346 (дата звернення: 21.10.2025). DOI: 10.

10. Michael is better than Mehmet: exploring the perils of algorithmic biases and selective adherence to advice from automated decision support systems in hiring. Frontiers in Psychology, 2024. 2024.1416504/full. 2024.1416504 URL: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2024.1416504/full (дата звернення: 21.10.2025). DOI: 10.3389/fpsyg.

11. Munoriyarwa A., de-Lima-Santos M.-F. Generative AI and the Future of News: Examining AI's Agency, Power, and Authority. *Digital Journalism*, 2025, 1–15. 2025.2545448. URL: 1080/17512786.2025.2545448. 2025.2545448 (дата звернення: 21.10.2025)..

12. Padalko O. AI-Driven Disinformation: Policy Recommendations for Democratic Resilience. Frontiers in Artificial Intelligence, 2025, 7, 1–12. 2024.1415908. URL: 3389/frai.2024.1415908. (дата звернення: 21.10.2025).

13. Royal United Services Institute (RUSI). Russia, AI and the Future of Disinformation Warfare. 2024. URL: https://static.rusi.org/russia-ai-and-the-future-of-disinformation-warfare.pdf (дата звернення: 21.10.2025)

14. Shen C. Fair Use, Licensing, and Authors' Rights in the Age of Generative AI. Northwestern. *Journal of Technology and Intellectual Property*, 2024, 22(1), 157–180. URL: https://scholarlycommons.law.northwestern.edu/njtip/vol22/iss1/4 (дата звернення: 21.10.2025)

15. Su D., та ін. RAG-Fusion Based Information Retrieval for Fact-Checking. arXiv preprint arXiv:2406.18866, 2024. 2406.18866. 2406.18866 URL: 48550/arXiv.2406.18866 (дата звернення: 21.10.2025). DOI: 10.

16. Taeihagh A. Governance of artificial intelligence. Policy and Society, 2024, 44(1), 1–28. URL: 1093/polsoc/puad023. (дата звернення: 21.10.2025). DOI: 10.

17. The Evolving Role of AI in Russian Disinformation Campaigns. Ukraine Analytica, 2024, 3(35), 9–11. URL: https://ukraine-analytica.org/wp-content/uploads/UA_Analytica_35_2024.pdf (дата звернення: 21.10.2025)

18. Thomson A., та ін. AI and the future of visual journalism: Interrogating the threats of misinformation, objectivity, and labour displacement. Digital Journalism, 2025, 1–20. 2025.2451677. URL: 1080/17512786.2025.2451677. 2025.2451677 (дата звернення: 21.10.2025). DOI: 10.

19. Toff B., Simon F. M. 'It's a question of where the human is': Audience perceptions of AI in news production. Digital Journalism, 2024, 1–23. 2024.2398642. 2024.2398642 URL: 1080/21670811.2024.2398642. 2024.2398642 (дата звернення: 21.10.2025). DOI: 10.

20. Torrance A. W., Tomlinson B. Training is Everything: Artificial Intelligence, Copyright, and "Fair Training". Dickinson Law Review, 2024, 128(3), 637–662. URL: https://insight.dickinsonlaw.psu.edu/dlr/vol128/iss3/3 (дата звернення: 21.10.2025)

21. World Privacy Forum. Privacy, Identity and Trust in C2PA: A Technical Review and Analysis of the C2PA Digital Media Provenance Framework. 2025. URL: https://worldprivacyforum.org/posts/privacy-identity-and-trust-in-c2pa/ (дата звернення: 21.10.2025)

Kanishchev S. PROBLEMS OF AI CONTENT GENERATION IN JOURNALISM: EPISTEMIC, ETHICAL, AND ORGANIZATIONAL DIMENSIONS

This article proposes and substantiates a comprehensive AI Governance System for Newsrooms (AGSN) – a new, process-oriented model designed to minimize the epistemic, ethical, and organizational risks associated with the integration of generative artificial intelligence (AI) into journalistic practice. The relevance of the study is driven by the rapid yet often unsystematic adoption of generative models in editorial processes, which poses threats to the credibility, transparency, and accountability of the media. The purpose of this article is to develop a practical toolkit that enables newsrooms to structure their work with AI while upholding core journalistic standards. The research methodology is based on a qualitative analysis and synthesis of existing editorial policies from leading global media outlets, a review of current academic publications on AI risks, and the application of process management and risk management principles to the journalistic production cycle. The results of the work are presented in the form of the AGSN, which includes several key components: 1) a detailed operational protocol: "prompt, generation, verification, labeling, publication, post-publication monitoring". 2) a risk assessment matrix for various journalistic genres. 3) recommendations for implementing technological solutions, such as text generation grounded in verified knowledge bases to reduce factual errors and the use of standards for verifying media content provenance. 4) a set of editorial policies on labeling, fact-checking, and defining "red lines" for AI use. Special attention is given to the application of the system in the Ukrainian context, where the conditions of information warfare place heightened demands on information security and countering disinformation. The scientific novelty lies in the transition from fragmented ethical recommendations to a holistic, operational governance system that combines technological, organizational, and ethical safeguards. The practical value of the work consists in providing newsrooms with a ready-to-implement model that standardizes work with generative AI, enhances risk manageability, and helps maintain audience trust.

Key words: *generative AI, journalism, ethics, fact-checking, disinformation, editorial policies.*

Дата надходження статті: 15.10.2025
Дата прийняття статті: 10.11.2025
Опубліковано: 29.12.2025