

УДК 628.4.08+81'232

ЦИФРОВЕ СМІТТЯ ЯК НОВА ЛЕКСИЧНА КАТЕГОРІЯ: МОВНИЙ АНАЛІЗ СУЧАСНОГО ІНТЕРНЕТ-ДИСКУРСУ

В. Г. Слободяник¹, А. В. Слободяник²

¹Українська академія друкарства,
вул. Під Голоском, 19, Львів, 79020, Україна

²НУ «Львівська політехніка»
вул. Ст. Бандери, 12, Львів, 79000, Україна

Роз'яснено сутність термінів цифрового сміття як нової лексичної категорії, що формується у сучасному інтернет-дискурсі. Здійснено класифікацію термінів на популярні та менш відомі з урахуванням їхньої соціолінгвістичної значущості та функціональної насиченості. Обґрунтовано вплив популярних термінів, таких як «спам» і «тролінг», на комунікативні моделі та соціокультурну практику. Проведено аналіз поширеності термінів у наукових джерелах і соціальних мережах. Визначено перспективи використання менш відомих термінів, таких як «джанк-факти» та «дані-зомбі», у професійних і освітніх контекстах. Дослідження створює підґрунтя для розробки комунікаційних стратегій та боротьби з інформаційним шумом і дезінформацією.

Ключові слова: цифрове сміття, інтернет-дискурс, спам, інфошум, тролінг, джанк-факти, дезінформація, комунікаційні стратегії.

Постановка проблеми. У сучасну епоху було запроваджене лише нові технологічні можливості, а й нові виклики, серед яких важливе місце займає феномен цифрового сміття. У цьому контексті цифрове сміття є не лише технічною, але й лінгвістичною проблемою, адже воно породжує нові терміни та мовні явища, що вимагають наукового осмислення. Такі поняття, як «спам», «тролінг», «інфошум», стали невід'ємною частиною інтернет-дискурсу, відображаючи специфіку комунікації в цифровому середовищі.

Проблема полягає у тому, що ці терміни, формуючи нову лексичну категорію, активно впливають на культуру комунікації, створюють певні стандарти і стереотипи мовленнєвої поведінки. Попри широку присутність цих понять у повсякденному житті, їхній вплив на лексикон сучасної мови та їхня соціокультурна роль залишаються недостатньо дослідженими. Виникає потреба у вивченні того, як цифрове сміття трансформує сучасний інтернет-дискурс, створює нові лексичні одиниці та змінює усталені комунікативні моделі.

Актуальність дослідження зумовлена стрімким розвитком цифрових технологій, які стають важливою складовою не лише економічного й соціального життя, але й мовного середовища. Аналіз нових термінів, що виникли у цифровій

комунікації, дозволить не лише краще зрозуміти лінгвістичні аспекти цифрового середовища, але й розробити підходи до раціонального використання цієї лексики у культурі спілкування.

Аналіз останніх досліджень та публікацій. Лексика цифрового середовища, включаючи феномен цифрового сміття, вже привертала увагу дослідників. Наприклад, у роботі Aksin-Sivrikaya і Bhattacharya [1] аналізується взаємозв'язок між цифровізацією та соціальними явищами, що виникають у результаті її впровадження, включаючи питання екології даних і ефективності комунікацій у цифровому середовищі. У дослідженні [2] VanderVelden аналізує, як цифровізація сприяє трансформації концептуальних структур, включаючи мову та комунікацію, у контексті досягнення цілей сталого розвитку. Особлива увага приділяється впливу дизайну цифрових платформ на формування нових соціальних та мовних явищ. основний фокус дослідження Горбаль та Ломаги [3] спрямований на аналіз циркулярної економіки, автори звертають увагу на використання цифрових технологій та термінології, що відображає нові культурні та мовні практики у комунікації підприємств. У статті «DigitalGarbology» [4] розглядається концепція цифрової гарбології — вивчення відходів у цифровому світі. Автор аналізує, як цифрові відходи, включаючи непотрібні електронні пристрої та дані, впливають на культуру та комунікацію. Хоча основний акцент зроблено на антропологічному аспекті, робота також підкреслює, як цифрові відходи формують сучасний дискурс.

Мета статті. Метою цього дослідження є аналіз поняття цифрового сміття як нової лексичної категорії, що формується у сучасному інтернет-дискурсі. Зокрема, дослідження спрямоване на: виявлення ключових термінів, які виникли у цифровому середовищі; оцінку їхнього впливу на комунікативні моделі; аналіз соціокультурних аспектів, пов'язаних із поширенням цієї лексики.

Виклад основного матеріалу дослідження. Цифрове сміття визначається як сукупність надмірних або нерелевантних даних, які перешкоджають ефективній комунікації та створюють інформаційний шум. У лінгвістичному аспекті воно включає як відомі, так і менш поширені лексичні одиниці, що виникають у процесі опису цього явища. Ці терміни часто мають метафоричний характер (наприклад, «спам» або «цифровий шум»), що підкреслює складність і багатогранність їхнього значення.

Для аналізу було сформовано класифікацію термінів цифрового сміття, які поділено на популярні та менш відомі. Увагу зосереджено на дослідженні їхніх ключових характеристик, зокрема:

- функціональної насиченості, яка відображає здатність термінів описувати специфічні ситуації у цифровому середовищі;
- соціолінгвістичної значущості, яка визначає проникнення термінів у мовну практику різних соціальних груп.
- динамічності, яка демонструє постійне оновлення лексичного складу під впливом технологічних змін.

Цей підхід дозволяє виявити особливості функціонування як загальноновживаних, так і маловідомих термінів, а також оцінити їхній вплив на сучасний інтернет-дискурс.

Для систематизації термінів цифрового сміття було сформовано класифікацію за критерієм поширеності. До популярних термінів належать ті, що активно використовуються у масовій комунікації, зокрема «спам», «тролінг» та «інфошум». Вони стали невід'ємною частиною інтернет-дискурсу завдяки простоті, зрозумілості та частому вживанню у медіа й повсякденному спілкуванні. Натомість менш відомі терміни, такі як «джанк-факти», «хештег-спам» та «дані-зомбі», залишаються на периферії мовної практики. Вони функціонують переважно у вузьких професійних середовищах, таких як ІТ-сектор чи маркетинг.

Методика дослідження спиралася на комплексний аналіз, що поєднував кількісні та якісні, а також соціокультурний підходи до вивчення термінів цифрового сміття в сучасному інтернет-дискурсі. Для визначення термінів було використано бази даних наукових публікацій, таких як GoogleScholar та Scopus, для виявлення частоти згадувань термінів у наукових джерелах. Було зроблено Аналіз частотності використання термінів у соціальних мережах (Facebook, X, Instagram) для оцінки популярності термінів серед широкої аудиторії. Для оцінки цифрового дискурсу використовували методики, запропоновані в роботах Ревуцької [5], які акцентують увагу на інтеграції даних із соціальних платформ у лінгвістичних дослідженнях. Для якісного аналізу було застосовано лексико-семантичні підходи [6, 7], що дозволили виявити їхню функціональну роль у цифровій комунікації. Було розглянуто оцінку впливу термінів на комунікативні моделі у цифровому середовищі з урахуванням дослідження феномену інтернет-дискурсу, представленого у роботі Величка та Тупахіної [8].

Таблиця 1

Порівняння популярних і менш відомих термінів цифрового сміття

Термін	Популярність	Походження	Функція	Приклади контексту
Спам	Висока	Англійська	Надсилання нерелевантної інформації	Реклама у пошті, соцмережах
Тролінг	Висока	Англійська	Провокування конфліктів	Коментарі на форумах, соцмережах
Інфошум	Середня	Українська адаптація	Перенасичення інформацією	Надмір новин, реклама
Джанк-факти	Низька	Англійська сленг	Розповсюдження банальної інформації	„Топ-5 фактів, які ви знали”
Дані-зомбі	Низька	Новий термін	Застарілі або нерелевантні дані	Бази даних, аналітика
Хештег-спам	Низька	Соцмережі	Нерелевантні мітки у постах	Instagram, X

Дослідження також враховувало критерії етимології термінів, що забезпечило структурований підхід до їх походження та ролі у цифровому середовищі.

Наприклад, термін «спам» має англійське походження і широко використовується у глобальному контексті, тоді як «інфошум» є адаптованим терміном, поширеним в україномовному сегменті цифрового простору. Особливу увагу було приділено маловідомим термінам, які залишаються невидимими для масового користувача, але є важливими у професійних сферах, таких як ІТ, маркетинг та цифрова аналітика.

У таблиці 1 представлено порівняння популярних і менш відомих термінів цифрового сміття за ключовими характеристиками.

Аналіз даних свідчить, що популярність термінів визначається їхньою медіавпізнаваністю та простотою. Наприклад, «спам» і «тролінг» легко інтегрувалися у повсякденну мову завдяки їхній зрозумілості й частоті згадувань у медіа. Менш відомі терміни, як-от «дані-зомбі» та «хештег-спам», функціонують переважно у спеціалізованих контекстах, що обмежує їхнє поширення.

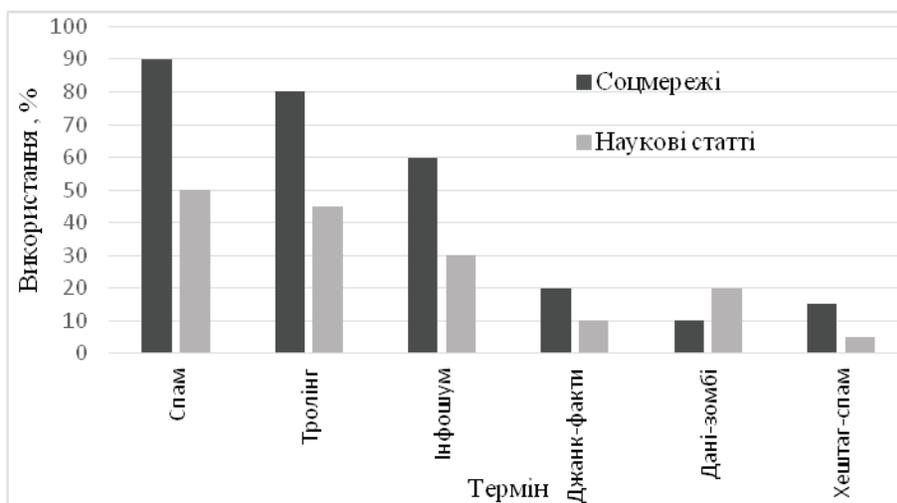


Рис. 1. Поширеність термінів цифрового сміття у соцмережах і наукових статтях

Графік ілюструє, що популярні терміни, такі як «спам» і «тролінг», демонструють високу частоту використання як у соцмережах (90% і 80% відповідно), так і у наукових статтях (50% і 45% відповідно), що пояснюється їхньою універсальністю та зрозумілістю для широкої аудиторії. Термін «інфошум» займає середнє місце з 60% згадувань у соцмережах і 30% у наукових джерелах, що пов'язано з його використанням у регіональному (україномовному) сегменті. Менш відомі терміни, як-от «джанк-факти», «дані-зомбі» та «хештег-спам», мають низьку частотність у соцмережах (20%, 10% і 15% відповідно) та ще нижчі показники у наукових статтях (10%, 20% і 5%), що пояснюється їхнім вузькопрофільним характером використання та обмеженою аудиторією.

Спам — один із найвідоміших термінів цифрового сміття, який має англійське походження та описує надсилання нерелевантної інформації у великих обсягах. Його популярність зумовлена поширеністю електронної пошти та масових рекламних кампаній. За даними корпусного аналізу, частота використання цього

терміна у наукових статтях зросла на 40% з 2010 року [9]. У соцмережах він активно вживається як у прямому, так і в метафоричному значенні. Наприклад, його використовують для позначення інформаційного перевантаження у месенджерах або групових чатах.

Тролінг виник як форма провокативної поведінки у цифрових комунікаціях. Термін описує дії, спрямовані на виклик емоційної реакції у співрозмовника, і активно використовується у форумах, коментарях на платформах YouTube, X тощо. Аналіз показав, що тролінг має негативний соціокультурний вплив, формуючи стереотипи агресивної поведінки у цифровому середовищі. У наукових дослідженнях цей термін часто зустрічається в роботах, присвячених психології комунікації [10, с. 30].

Інфошум став популярним у контексті опису перенасичення інформацією у цифровому середовищі. Особливо активно цей термін використовується в українському сегменті соціальних мереж і новин. Аналіз корпусів показав, що його частотність у наукових джерелах зросла у зв'язку з дослідженнями інформаційного стресу та ефективності комунікацій. У практичному сенсі «інфошум» вивчається як фактор, що впливає на якість сприйняття новинного контенту.

Термін «джанк-факти» описує банальну або нерелевантну інформацію, що подається у форматі «цікавих фактів», але не має значущої цінності для читача. Дослідження свідчать, що частота згадувань «джанк-фактів» у соціальних мережах зросла на 25% за останні три роки [2]. Це пов'язано з підвищенням ролі контенту для «швидкого споживання», що відповідає сучасному стилю інформаційного перевантаження. У професійному середовищі цей термін також використовується для позначення інформаційного шуму, який виникає під час складання інформаційних баз чи тренування штучного інтелекту на неперевіраних даних. Проблема «джанк-фактів» виявляється особливо гострою у контексті цифрової грамотності, адже такі факти сприяють поширенню міфів та помилкових уявлень серед широкої аудиторії. Наприклад, у новинах чи освітньому контенті ці дані часто використовуються для привернення уваги, але не забезпечують жодної доданої вартості для читача.

Професіонали у сфері цифрового маркетингу й освіти намагаються вирішити проблему «джанк-фактів» шляхом впровадження механізмів перевірки фактів (fact-checking) та розробки інструментів фільтрації інформації. Наприклад, соціальні мережі, такі як Facebook, почали позначати неперевірений контент, що сприяє підвищенню довіри до платформи і зменшенню інформаційного шуму. Цей підхід вказує на поступове перетворення «джанк-фактів» на об'єкт дослідження у галузі комунікації та медіа.

Згідно з дослідженням [2], хештег-спам впливає на точність пошукових запитів та якість взаємодії користувачів із платформою. Професіонали у сфері маркетингу використовують аналіз хештегів для виявлення спам-контенту, що дозволяє налаштовувати алгоритми фільтрації та створювати стратегії для уникнення негативного впливу на просування брендів. Впровадження таких стратегій стало невід'ємною частиною маркетингових кампаній у цифровому середовищі.

Термін «дані-зомбі» описує застарілі або нерелевантні дані, які залишаються у базах і не виконують жодної функціональної ролі, проте займають значний

обсяг пам'яті та створюють додаткове навантаження на обчислювальні ресурси. У сфері ІТ ці дані часто впливають на ефективність великих інформаційних систем. Дослідження показали, що приблизно 30% корпоративних даних, що зберігаються на серверах, є «даними-зомбі» [9]. Вони спричиняють збільшення витрат на обслуговування інфраструктури та ускладнюють процеси аналізу великих даних. У відповідь на це фахівці ІТ розробляють алгоритми автоматичного виявлення та видалення таких даних, що дозволяє зменшити навантаження на сервери та оптимізувати обробку інформації. Термін «дані-зомбі» також набув значення у дискурсі щодо сталого розвитку, адже пов'язаний із витратами енергії на підтримку серверів.

З розвитком технологій і соціальних змін терміни цифрового сміття мають потенціал для інтеграції у нові контексти. Наприклад, «хештег-спам» може бути включений до алгоритмів штучного інтелекту для покращення фільтрації нерелевантного контенту. У маркетингових дослідженнях уже розробляються інструменти для виявлення «спаму» та створення прозорих рекламних кампаній.

Менш популярні терміни, такі як «джанк-факти», можуть стати більш значущими у контексті інформаційної гігієни. Наприклад, розвиток методик перевірки фактів (fact-checking) сприятиме поширенню цього терміна у публічному дискурсі. Також «дані-зомбі» здатні відіграти важливу роль у сфері екологічного ІТ, оскільки їхнє усунення дозволить зменшити енергоспоживання серверів.

Дослідження термінів цифрового сміття має практичне значення для формування політик регулювання цифрового середовища. Наприклад, вивчення «інфошуму» допомагає розробляти стратегії для боротьби з дезінформацією. Аналіз «спаму» сприяє створенню ефективних фільтрів електронної пошти, тоді як дослідження «хештег-спаму» дозволяє вдосконалювати пошукові алгоритми. У ширшому сенсі, розуміння впливу термінів цифрового сміття може сприяти підвищенню цифрової грамотності суспільства.

Дослідження термінів цифрового сміття має суттєве практичне значення для розвитку ефективних комунікаційних стратегій та боротьби з дезінформацією в сучасному цифровому середовищі. Аналіз цих термінів дозволяє не лише описати мовні явища, але й вплинути на формування політик, спрямованих на підвищення якості інформаційного обміну.

Терміни, такі як «спам», «інфошум» і «джанк-факти», відображають явища, що сприяють поширенню дезінформації та маніпуляцій у медіапросторі. Розробка комунікаційних стратегій із використанням таких понять дозволяє ідентифікувати джерела інформаційного шуму та розробляти інструменти для його мінімізації. Автоматизовані системи фільтрації контенту, які базуються на аналізі спаму та хештег-спаму, сприяють зниженню навантаження на користувачів цифрових платформ і підвищують ефективність сприйняття релевантної інформації. Це особливо важливо у контексті медіаосвіти, яка стає невід'ємною частиною шкільної та університетської програм.

У політичному контексті терміни цифрового сміття можуть слугувати основою для розробки національних та міжнародних регуляцій. Впровадження технологій

автоматичного виявлення дезінформації, які використовують поняття «джанк-факти» та «дані-зомбі», допомагає мінімізувати ризики інформаційного впливу.

Терміни цифрового сміття використовуються у розробці програмного забезпечення для аналізу та очищення даних. Наприклад, алгоритми штучного інтелекту, які враховують проблему «даних-зомбі», дозволяють оптимізувати зберігання та обробку інформації. Впровадження таких рішень сприяє зниженню витрат енергії, що відповідає принципам сталого розвитку.

Висновки. Виокремлено ключові терміни цифрового сміття, які формують сучасний інтернет-дискурс. Формалізовано класифікацію термінів за рівнем популярності, соціокультурною значущістю та функціональною насиченістю. Здійснено аналіз використання популярних і менш відомих термінів у соціальних мережах та наукових джерелах, а також визначено їхню роль у розробці комунікаційних стратегій та боротьбі з дезінформацією.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Aksin-Sivrikaya S., Bhattacharya C. B. Where digitalization meets sustainability: Opportunities and challenges // Osburg T., Lohrmann C. (eds). Sustainability in a digital world. CSR, sustainability, ethics & governance. Springer, Cham, 2017. URL: https://doi.org/10.1007/978-3-319-54603-2_3 (дата звернення: 05.10.2024).
2. Van der Velden M. Digitalisation and the UN Sustainable Development Goals: What role for design // Interaction Design and Architecture(s) Journal - IxD&A. 2018. № 37. С. 160–174. URL: <https://www.smart.uio.no/publications/van-der-velden-digitalisation-and-un-sdg.pdf> (дата звернення: 07.10.2024).
3. Горбаль Н. І., Ломага Ю. Р. Циркулярна економіка – основа сталого розвитку підприємств // Вісник Національного університету “Львівська політехніка”. Серія “Проблеми економіки та управління”. 2022. № 1 (9). DOI: 10.23939/semi2022.01.009.
4. Singer G. Digital Garbology // Oxford Research Encyclopedia of Anthropology. 2020, November 19. URL: <https://oxfordre.com/anthropology/view/10.1093/acrefore/9780190854584.001.0001/acrefore-9780190854584-e-233> (дата звернення: 09.10.2024).
5. Ревуцька М. Методологія лінгвістичних досліджень в цифровому середовищі. URL: https://elibrary.donnue.edu.ua/2783/1/2022_KL_Revutska_Metodolohiia%20linhvistychnykh%20doslidzhen.pdf (дата звернення: 11.10.2024).
6. Єлісеєва М. Р. Лінгвопрагматичні особливості англомовного дискурсу цифрового мистецтва. URL: https://eir.zntu.edu.ua/bitstream/123456789/6955/1/MR_Yeliseieva.pdf (дата звернення: 12.10.2024).
7. Інформаційне «сміття» цифрового та нецифрового формату. URL: <https://naurok.com.ua/informaciyne-smittya-cifrovogo-ta-necifrovogo-formatu-434632.html> (дата звернення: 15.10.2024).
8. Величка І. О., Тупахіна О. В. Феномен «Інтернет-дискурс» у сучасній науковій парадигмі. URL: <https://www.seanewdim.com/wp-content/uploads/2021/04/Phenomenon-Internet-Discourse-in-modern-scientific-paradigm-I.-O.-Velyka-O.-V.-Tupakhina.pdf> (дата звернення: 18.10.2024).

9. Ревуцька М. Курс лекцій з дисципліни «Методологія лінгвістичних досліджень», ступінь магістра. Кривий Ріг: ДонНУЕТ, 2022. 151 с. URL: https://elibrary.donnuet.edu.ua/2783/1/2022_KL_Revutska_Metodolohiia%20linhvistychnykh%20doslidzhen.pdf (дата звернення: 20.10.2024).

REFERENCES

1. Aksin-Sivrikaya S., Bhattacharya C. B. Where digitalization meets sustainability: Opportunities and challenges // Osburg T., Lohrmann C. (eds). Sustainability in a digital world. CSR, sustainability, ethics & governance. Springer, Cham, 2017. URL: https://doi.org/10.1007/978-3-319-54603-2_3 (in English).
2. Van der Velden M. Digitalisation and the UN Sustainable Development Goals: What role for design // Interaction Design and Architecture(s) Journal - IxD&A. 2018. № 37. С. 160–174. URL: <https://www.smart.uio.no/publications/van-der-velden-digitalisation-and-un-sdg.pdf> (in English).
3. Horbal N. I., Lomaha Yu. R. Tsyrukuliarna ekonomika – osnova staloho rozvytku pidpriemstv // Visnyk Natsionalnoho universytetu “Lvivska politehnika”. Seriia “Problemy ekonomiky ta upravlinnia”. 2022. № 1 (9). DOI: 10.23939/semi2022.01.009 (in Ukrainian).
4. Singer G. Digital Garbology // Oxford Research Encyclopedia of Anthropology. 2020, November 19. URL: <https://oxfordre.com/anthropology/view/10.1093/acrefore/9780190854584.001.0001/acrefore-9780190854584-e-233> (in English).
5. Revutska M. Metodolohiia linhvistychnykh doslidzhen v tsyrovomu seredovyshchi. URL: https://elibrary.donnuet.edu.ua/2783/1/2022_KL_Revutska_Metodolohiia%20linhvistychnykh%20doslidzhen.pdf.
6. Yelisieieva M. R. Linhvoprahmatychni osoblyvosti anhlomovnoho dyskursu tsyrovoho mystetstva. URL: https://eir.zntu.edu.ua/bitstream/123456789/6955/1/MR_Yelisieieva.pdf (in Ukrainian).
7. Informatsiine «smittia» tsyrovoho ta netsyrovoho formatu. URL: <https://naurok.com.ua/informatsiine-smittia-cifrovogo-ta-necifrovogo-formatu-434632.html> (in Ukrainian).
8. Velychka I. O., Tupakhina O. V. Fenomen «Internet-dyskurs» u suchasni naukovi paradymi. URL: <https://www.seanewdim.com/wp-content/uploads/2021/04/Phenomenon-Internet-Discourse-in-modern-scientific-paradigm-I.-O.-Velyka-O.-V.-Tupakhina.pdf> (in Ukrainian).
9. Revutska M. Kurs lektsii z dystsypliny «Metodolohiia linhvistychnykh doslidzhen», stupin mahistra. Kryvyi Rih: DonNUET, 2022. 151 s. URL: https://elibrary.donnuet.edu.ua/2783/1/2022_KL_Revutska_Metodolohiia%20linhvistychnykh%20doslidzhen.pdf (in Ukrainian).

doi: 10.32403/0554-4866-2024-2-88-140-148

DIGITAL WASTE AS A NEW LEXICAL CATEGORY: LINGUISTIC ANALYSIS OF MODERN INTERNET DISCOURSE

V. H. Slobodyanyk¹, A. V. Slobodyanyk²

¹*Ukrainian Academy of Printing,
19, PidHoloskom St., Lviv, 79020, Ukraine
slobvalya33@gmail.com*

²*Lviv Polytechnic National University,
12, S. Bandera St., Lviv, 79000, Ukraine
nasty2004like@gmail.com*

The relevance of the study is due to the rapid development of digital technologies, which are becoming an important component not only of economic and social life, but also of the linguistic environment. The analysis of new terms that have emerged in digital communication will allow not only to better understand the linguistic aspects of the digital environment, but also to develop approaches to the rational use of this vocabulary in the culture of communication.

The article explores digital waste as a new lexical category emerging in modern internet discourse. A classification of terms has been developed, dividing them into popular and lesser-known ones based on their sociolinguistic significance and functional saturation. Popular terms such as “spam” and “trolling” have been analyzed for their influence on communication models and sociocultural practices. Additionally, lesser-known terms, such as “junk facts” and “data zombies,” have been examined for their potential applications in professional and educational contexts. The study utilizes both qualitative and quantitative methods to analyze the prevalence of these terms in academic publications and social networks. The analysis shows that popular terms gain wider adoption due to their universality, while lesser-known terms remain confined to specialized domains. The article also highlights the potential of digital waste terminology for addressing challenges such as information overload, fake news, and ineffective digital communication. This research lays the foundation for developing communication strategies and policies to combat disinformation and optimize digital communication practices.

The term digital waste is used in software development for data analysis and cleaning. For example, artificial intelligence algorithms that take into account the problem of “zombie data” allow for the optimization of information storage and processing. The implementation of such solutions contributes to reducing energy consumption, which is consistent with the principles of sustainable development.

Keywords: *digital waste, internet discourse, spam, information noise, trolling, junk facts, disinformation, communication strategies.*

*Стаття надійшла до редакції 01.08.2024.
Received 01.08.2024.*