

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ
ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ
КАФЕДРА ІНЖЕНЕРІЇ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ**

КВАЛІФІКАЦІЙНА РОБОТА
на тему: «Методика реферування новинних статей на
основі метаданих»

на здобуття освітнього ступеня магістра
зі спеціальності 121 Інженерія програмного забезпечення
освітньо-професійної програми Інженерія програмного забезпечення

*Кваліфікаційна робота містить результати власних досліджень.
Використання ідей, результатів і текстів інших авторів мають посилання на
відповідне джерело.*

_____ Анна ГОНЧАРОВА
(підпис)

Виконав: здобувач вищої освіти група ПДМ-62

Анна ГОНЧАРОВА

Керівник: Оксана ЗОЛОТУХІНА
к.т.н., доцент

Рецензент:

Київ 2025

**ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ**
Навчально-науковий інститут інформаційних технологій

Кафедра Інженерії програмного забезпечення

Ступінь вищої освіти Магістр

Спеціальність 121 Інженерія програмного забезпечення

Освітньо-професійна програма «Інженерія програмного забезпечення»

ЗАТВЕРДЖУЮ
Завідувач кафедри
Інженерії програмного забезпечення

_____ Ірина ЗАМРІЙ
« _____ » _____ 2024 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

Гончаровій Анні Андріївні

1. Тема кваліфікаційної роботи: Методика реферування новинних статей на основі метаданих
керівник кваліфікаційної роботи Оксана ЗОЛОТУХІНА к.т.н., доцент,
затверджені наказом Державного університету інформаційно-комунікаційних технологій від «15» жовтня 2024 р. № 320.
2. Строк подання кваліфікаційної роботи «17» грудня 2024 р.
3. Вихідні дані до кваліфікаційної роботи: науково-технічна література, опис методів обробки текстових даних, опис моделей та методів реферування, види метаданих новинних статей, вимоги до реферату.
4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)
 1. Дослідження методів автоматичного реферування новинних статей.
 2. Порівняльний аналіз технологій та переваг, недоліків та ефективності методів автоматизованого реферування.

3. Розробка методу обробки та систематизації новинного контенту за рахунок автоматизованого реферування новинних статей на основі метаданих.
5. Перелік ілюстративного матеріалу: *презентація*
 1. Мета, об'єкт та предмет дослідження
 2. Актуальність роботи
 3. Підходи до реферування
 4. Математична модель реферування
 5. Параметри семантичного аналізу новинних текстів
 6. Модифікований метод
 7. Структура системи автоматизованого реферування
 8. Практичний результат
 9. Порівняння часу створення рефератів вручну та за допомогою методики
 10. Порівняння якості рефератів
 11. Висновки
 12. Публікації та апробація роботи
6. Дата видачі завдання «16» жовтня 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Аналіз наявної науково-технічної літератури	16.10-22.10.2024	
2	Дослідження методів реферування	23.10-30.10.2024	
3	Проведення порівняльного аналізу ефективності існуючих методик реферування	01.11-07.11.2024	
4	Визначення напрямків удосконалення	08.11.2024	
5	Розробка архітектури методики автоматизованого реферування	09.11-19.11.2024	
6	Оформлення роботи: вступ, висновки, реферат	20.11-22.11.2024	
7	Розробка демонстраційних матеріалів	23.11-25.11.2024	
8	Попередній захист роботи	26.11- 06.12.2024	

Здобувач вищої освіти

_____ (підпис)

Анна ГОНЧАРОВА

Керівник кваліфікаційної роботи

_____ (підпис)

Оксана ЗОЛОТУХІНА

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття освітнього ступеня магістра: 84 стор., 9 табл., 2 рис., 35 джерел.

Мета дослідження – підвищення якості обробки та систематизації новинного контенту за рахунок автоматизованого реферування новинних статей на основі метаданих.

Об'єкт дослідження – процес автоматизованого реферування новинних статей.

Предмет дослідження – методи, моделі та алгоритми автоматизованого реферування новинних статей.

У роботі використано різноманітні методи, такі як структурно-семантичний аналіз текстів, моделювання за допомогою нейронних мереж, порівняльний аналіз існуючих підходів до автоматизації реферування, а також статистичні методи оцінки точності та релевантності рефератів.

Проведено аналіз сучасних методів реферування текстів у новинній журналістиці, включаючи екстрактивні, абстрактивні та гібридні підходи, а також їхню ефективність і застосування для створення якісних рефератів.

Розроблено та оптимізовано методику автоматизованого реферування новинних статей на основі метаданих, що дозволяє підвищити точність і швидкість обробки текстів.

Проведено експерименти для оцінки ефективності запропонованої методики, які підтвердили її переваги в порівнянні з традиційними підходами. Таким, досягнуто підвищення точності визначення ключових змістових елементів текстів.

КЛЮЧОВІ СЛОВА: АВТОМАТИЗОВАНЕ РЕФЕРУВАННЯ, МЕТАДАННІ, НОВИННІ СТАТТІ, ЕКСТРАКТИВНІ МЕТОДИ, АБСТРАКТИВНІ МЕТОДИ, НЕЙРОННІ МЕРЕЖІ, СТРУКТУРНО-СЕМАНТИЧНИЙ АНАЛІЗ.

ABSTRACT

Text part of the master's qualification work: 84 pages, 9 tables, 2 pictures, 35 sources.

The purpose of the research is to improve the quality of processing and systematization of news content through automated summarization of news articles based on metadata.

The object of the research is the process of automated summarization of news articles.

The subject of the research is methods, models and algorithms of automated summarization of news articles.

The work uses various methods, such as structural and semantic analysis of texts, modeling using neural networks, comparative analysis of existing approaches to automating summarization, as well as statistical methods for assessing the accuracy and relevance of summaries.

An analysis of modern methods of summarizing texts in news journalism is conducted, including extractive, abstract and hybrid approaches, as well as their effectiveness and application for creating high-quality summaries.

A method of automated summarization of news articles based on metadata has been developed and optimized, which allows to increase the accuracy and speed of text processing.

Experiments were conducted to assess the effectiveness of the proposed method, which confirmed its advantages in comparison with traditional approaches. Thus, an increase in the accuracy of determining key content elements of texts was achieved.

KEYWORDS: AUTOMATED SUBMITTING, METADATA, NEWS ARTICLES, EXTRACTIVE METHODS, ABSTRACTIVE METHODS, NEURAL NETWORKS, STRUCTURAL-SEMANTIC ANALYSIS.

ЗМІСТ

ВСТУП	9
1 ІНФОРМАЦІЙНО-АНАЛІТИЧНА ДІЯЛЬНІСТЬ. АНАЛІЗ ПРОБЛЕМИ ТА ОСНОВИ ВИКОРИСТАННЯ МЕТАДАНИХ НА ВЕБ-РЕСУРСАХ	13
1.1 Огляд інформаційно-аналітичної діяльності в контексті покращення процесу пошуку інформації на веб-ресурсах	13
1.2 Сучасний стан реферування текстів у новинній журналістиці	16
1.3 Порівняння з альтернативними методами	25
2 РОЗРОБКА МЕТОДИКИ РЕФЕРУВАННЯ НОВИННИХ СТАТЕЙ	37
2.1. Архітектура методики реферування	37
2.2. Практична реалізація методики	44
3 ВПРОВАДЖЕННЯ МЕТОДИКИ ТА ПЕРСПЕКТИВИ РОЗВИТКУ	48
3.1. Інтеграція методики у редакційні процеси	48
3.2. Розвиток та адаптація методики	56
ВИСНОВКИ	67
ПЕРЕЛІК ПОСИЛАНЬ	71
ДОДАТОК А. ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ	75
ДОДАТОК Б. ЛІСТИНГИ ПРОГРАМНИХ МОДУЛІВ	82

ВСТУП

У сучасному світі інформаційних технологій журналістика відіграє ключову роль у формуванні суспільної думки та забезпеченні громадян актуальними новинами. Зростання обсягів інформації, що публікується щоденно, ускладнює швидке засвоєння новинного контенту аудиторією. У цьому контексті методики автоматизованого реферування текстів стають необхідним інструментом для обробки великих обсягів інформації.

Реферування новинних статей сприяє підвищенню доступності інформації, дозволяючи аудиторії швидко отримувати основну суть текстів без втрати ключових деталей. Це особливо важливо в умовах сучасної медіареволюції, коли увага читача стає все більш розпорошеною, а час на споживання контенту обмежений. Актуальність теми також зумовлена необхідністю підвищення якості новинного контенту. Метадані, як інструмент структуризації інформації, забезпечують більш точне виділення ключових ідей та сприяють ефективному створенню рефератів. Це дозволяє зменшити ризик втрати важливих деталей і забезпечує краще розуміння тексту.

Зростаюча потреба в автоматизації обробки новинного контенту зумовлена не лише його обсягом, але й швидкістю появи нової інформації. Ручна обробка та створення рефератів новин стає все менш ефективною, оскільки потребує значних часових та людських ресурсів. При цьому важливо зберігати баланс між швидкістю обробки та якістю отриманих результатів. Автоматизація процесу реферування дозволяє значно скоротити час на опрацювання матеріалів та зменшити навантаження на редакторів.

Інтеграція автоматизованих методик реферування у редакційні процеси відкриває нові можливості для оптимізації роботи журналістів. Замість витрачання часу на рутинні завдання, фахівці можуть зосередитися на створенні якісного контенту, залишаючи обробку текстів автоматизованим системам. Крім того, зростання використання штучного інтелекту у медіагалузі обґрунтовує

необхідність створення інтелектуальних інструментів для реферування текстів. Це дозволяє враховувати складні мовні та семантичні аспекти, які важливі для створення адекватних рефератів.

Особливу роль у процесі автоматичного реферування відіграють метадані статей. Метадані, такі як заголовки, підзаголовки, теги, категорії, часові мітки та ключові слова, містять важливу інформацію про структуру та зміст тексту. Використання метаданих дозволяє підвищити якість реферування, оскільки вони часто містять уже виділену автором або редактором ключову інформацію. Крім того, метадані допомагають краще зрозуміти контекст новини та її релевантність для різних категорій читачів.

Додатково варто зазначити, що в умовах інформаційної війни та зростання кількості фейкових новин, автоматизація процесу реферування також може допомогти у швидкій ідентифікації та категоризації потенційно недостовірної інформації, оскільки метадані часто містять важливі маркери для такого аналізу. Автоматизовані системи реферування, базовані на метаданих, дозволяють забезпечувати достовірність і прозорість новин, що є критично важливим для збереження довіри аудиторії до медіа. Також це робить розробку подібних систем особливо актуальною в контексті сучасних інформаційних викликів.

На ринку існує багато складних рішень для автоматичного реферування, які використовують методи машинного навчання та штучного інтелекту. Однак такі системи часто вимагають значних обчислювальних ресурсів, складні в налаштуванні та потребують великих наборів даних для навчання. Для малих проектів, локальних новинних агентств та індивідуальних розробників існує потреба в простіших рішеннях, які можна легко впровадити та адаптувати під конкретні потреби. Розробка методики реферування на основі метаданих дозволяє створити ефективний інструмент, який не потребує складної інфраструктури та може бути реалізований з використанням базових технологій.

Мета дослідження – підвищення якості обробки та систематизації новинного контенту за рахунок автоматизованого реферування новинних статей на основі метаданих.

Об'єкт дослідження – процес автоматизованого реферування новинних статей.

Предмет дослідження – методи, моделі та алгоритми автоматизованого реферування новинних статей.

Для досягнення мети вирішувалися наступні завдання:

1. Літературний огляд. Провести дослідження і аналіз існуючих наукових джерел, щоб отримати повне уявлення про різноманітні методи та технології, що використовуються у реферуванні, зокрема новинних статей.
2. Порівняльний аналіз. Проведено порівняльний аналіз ефективності існуючих методик реферування.
3. Моделювання. Розробити архітектуру методики автоматизованого реферування.
4. Програмування. Реалізувати прототип системи, що використовує метадані для створення рефератів.
5. Визначити критерії оцінки якості автоматизованих рефератів.
6. Дослідити можливості інтеграції методики у редакційні процеси та адаптації до різних форматів новин.

У ході роботи були застосовані сучасні методи наукового аналізу, зокрема структурно-семантичний аналіз текстів, моделювання за допомогою нейронних мереж, порівняльний аналіз існуючих підходів до автоматизації реферування [1], а також статистичні методи оцінювання точності та релевантності рефератів. Практична частина дослідження базувалася на використанні алгоритмів обробки тексту із застосуванням методів машинного навчання.

Практична значимість. Розроблена методика може бути інтегрована у редакційні процеси для підвищення ефективності та якості підготовки реферативних матеріалів. Впровадження цієї системи дозволяє значно скоротити час опрацювання великих обсягів текстових даних, що є особливо актуальним для сучасної журналістики та інформаційних агентств. Крім того, запропонована методика може бути використана для навчання студентів профільних спеціальностей, що працюють із текстовими даними, а також у бібліотечних і архівних установах.

1 ІНФОРМАЦІЙНО-АНАЛІТИЧНА ДІЯЛЬНІСТЬ. АНАЛІЗ ПРОБЛЕМИ ТА ОСНОВИ ВИКОРИСТАННЯ МЕТАДАНИХ НА ВЕБ-РЕСУРСАХ

1.1 Огляд інформаційно-аналітичної діяльності в контексті покращення процесу пошуку інформації на веб-ресурсах

Щоденно кожен пересічний громадянин стикається з великими обсягами інформації на різноманітних ресурсах, зокрема і новинних. Кожна людина має певну сферу інтересів і здійснює пошук інформації відповідно до своїх потреб. Часто людина зіштовхується з необхідністю ознайомитись з великою кількістю об'ємних матеріалів чи журналістських статей, але за браком часу, виокремлює лише частину інформації, яка не завжди повністю розкриває задану тему. З розвитком інформаційних технологій та зростанням обсягів даних, правильне структурування, опис та представлення документів та інформаційних ресурсів за допомогою метаданих стає все більш актуальним. Стандарти методів, такі як Дублінське ядро, забезпечують уніфікований підхід до цього процесу. Реферування та анотування як методи згортання інформації є необхідними інструментами для обробки великих обсягів даних для покращення процесу пошуку ключової інформації або конкретних статей за ключовими поняттями.

Інформаційно-аналітична діяльність – сукупність дій, що відбуваються в інтересах підвищення ефективності управлінської діяльності і мають на меті пізнання сутності, причин, тенденцій розвитку подій і явищ, розгляд і оцінювання ситуацій, вироблення на основі аналізу й обробки інформації висновків, рекомендацій, коментарів.

Основні задачі інформаційно-аналітичної діяльності:

- знайти в інформаційному потоці ті факти, що викликають інтерес;
- у сукупності виділених фактів знайти проблемну ситуацію, яку необхідно розв'язати;
- на підставі проведеного аналізу зробити висновки, внести пропозиції для прийняття рішень.

Аналітично-синтетична переробка документа здійснюється з метою отримання необхідної інформації, а також оцінки, зіставлення, узагальнення та подання отриманої інформації у формі, що відповідає запиту.

До інформаційно-аналітичної інфраструктури належать як найдавніші винаходи людства (бібліотеки, архіви, бібліографічні служби), так і сучасніші реферативні служби, органи НТІ, новітні фірми, що розробляють електронні (зокрема, мережеві) інформаційні ресурси. Інформаційно-аналітичні служби (підрозділи) функціонують у структурі органів влади, міністерствах і відомствах, системі НТІ, органах ЗМІ, сфері бізнесу, при політичних партіях і рухах тощо. Загальною відмітною особливістю цих служб є їх органічна інтегрованість у відповідні сфери діяльності, функціональний і організаційний симбіоз з їх соціальними інститутами і конкретними організаціями. У діяльності зазначених аналітичних служб, покликаних здійснювати інформаційний супровід керування у відповідних сферах, найчіткіше виявляється інформаційно-допоміжна природа інформаційної аналітики, що споріднює останню з іншими видами інформаційно-інфраструктурних, інформаційно-допоміжних служб, що традиційно впорядковують інформаційний простір, оптимізуючи і спрямовуючи рух інформаційних потоків, забезпечують збереження накопичених інформаційних ресурсів та ін [2].

Водночас інформаційна аналітика не є лише ще одним новим організаційним різновидом вищезазначених служб, її сутнісна, базова функція принципово відрізняється від тих завдань, що вирішують в інформаційному просторі ці служби. Ключовими серед цих завдань є насамперед: аналітичне перетворення інформації та її структурне впорядкування (інформаційний аналіз, узагальнення, згортання; консолідація великих інформаційних масивів у формі баз даних та їх поширення). Інформаційна аналітика, використовуючи всі можливості цих служб, активно оперуючи їх інформаційними продуктами і послугами, виконує, насамперед, завдання якісно-змістовного перетворення інформації, функціонально перетинаючись у цьому аспекті з науковою (виробництво нового знання) й управлінською (розробка варіантів рішень, сценаріїв) діяльністю [2].

Міжнародний стандарт ISO 15836-2003 («Information and documentation – The Dublin Core metadata element set») визначає універсальний набір метаданих для будь-яких інформаційних ресурсів. Стандарт покладений в основу практично всіх наборів метаданих, що описують текстові контенти. Будь-який документ характеризується контентом, тобто інформацією, в якій закріплено змістове навантаження та структурою, тобто зовнішнім виглядом та розміщенням частин контенту (наприклад, носій, формат, організація даних, розміщення реквізитів, шрифти, примітки тощо), а також наявністю у документі зв'язків з іншими документами (гіперпосилань). На базі Дублінського ядра було зроблено каталогізацію цифрових, де наведено нескладну систему опису інформаційних, зокрема, мережевих, ресурсів [3]. Для опису інформаційного ресурсу пропонується використовувати 29 характеристик, обов'язковими з яких є для всіх типів ресурсів:

1. Ідентифікатор інформаційного ресурсу (посилання на ресурс).
2. Назва ресурсу.
3. Власник.
4. Опис (зміст ресурсу, включаючи анотацію або реферат, або опис контенту візуальних, аудіо- або мультимедійних ресурсів);
5. Коди рубрикатора – тематика ресурсу, що виражена кодами стандартного переліку тематичних рубрик.
6. Ключові слова.
7. Мова (для текстового ресурсу).
8. Період оновлення.
9. Фінансування – форма фінансування при створенні та введенні ресурсу.
10. Дата останнього оновлення вмісту або дата створення.

Умовно обов'язкові:

1. Дата останнього оновлення метаданих інформаційного ресурсу.
2. Адреса в мережі.
3. Консультант – особа, до якої треба звертатися за додатковою інформацією.
4. Дата реєстрації інформаційного ресурсу;

5. Служба, що реєструє інформаційний ресурс.

Великий відсоток веб-контента генерується спеціальним програмним забезпеченням, наприклад, генераторами сайтів. Такі програмні засоби дозволяють генерувати пости, тобто вебконтент, фактично, дані, які супроводжуються метаданими. Кожний такий «генератор» може мати власні вимоги до метаданих та використовувати власні механізми створення метаданих та протоколи їх зберігання. Так, у Wintersmith обов'язковими вважають три види метаданих: `template` визначає шаблон, що використовується для рендерінга, `title` – назву поста, `date` – його дату. Але існує можливість додавання будь-яких метаданих. Метадані, як правило, мають власний формат представлення (окремий від формату даних). Одним з найбільш розповсюджених є YAML формат.

Широке використання YAML отримав завдяки тому, що він орієнтований на зручність введення/виведення типових структур даних багатьох мов програмування, тобто даних багатьох форматів, включаючи Json об'єкти. Він є дружнім до людини форматом серіалізації даних, концептуально близький до мов розмітки. Дозволяє визначити будь-які метадані відповідно до вимог бізнес – застосунку. Json об'єкти, в свою чергу, довели свою зручність та доцільність з розвитком великих даних та мережевих технологій, як стандартний текстовий формат для представлення структурованих даних на основі синтаксису об'єкта JavaScript, що, зазвичай, використовується для передачі даних до вебзастосунків [14].

1.2 Сучасний стан реферування текстів у новинній журналістиці

Реферування текстів у новинній журналістиці є складним і динамічним процесом, який постійно еволюціонує під впливом технологічних та інформаційних трансформацій сучасного медійного простору. Сьогодні це багатогранне явище, що має особливе значення для ефективної комунікації та швидкого поширення інформації в умовах інформаційного перенасичення.

Під реферуванням у контексті новинної журналістики розуміють специфічний аналітико-синтетичний процес трансформації первинного журналістського матеріалу, спрямований на створення гранично стислого, але інформативно місткого вторинного тексту, який максимально точно передає основний зміст та сенс оригінальної публікації. Це інтелектуальна діяльність, що вимагає глибокого розуміння контексту, структури та семантичних особливостей вихідного повідомлення. Сучасні методики реферування новинних текстів характеризуються комплексним підходом, що поєднує традиційні методи аналізу інформації з інноваційними технологіями штучного інтелекту та машинного навчання. Науковці та практики виділяють кілька принципових підходів до реферування: *extractive* (*extraction*), *abstractive* (*abstraction*) та гібридні методи, кожен з яких має власні переваги та обмеження.

Extractive-реферування передбачає пряме виділення та копіювання найбільш значущих фрагментів вихідного тексту без суттєвих змін їхньої структури та лексики. Цей метод є найбільш простим і технологічно доступним, однак має обмежену здатність до генерування дійсно якісних рефератів, оскільки механічно відтворює структуру оригінального тексту [1].

Abstractive-реферування є більш складним і наближеним до людського розуміння методом, який дозволяє не просто копіювати фрагменти тексту, а генерувати принципово новий текст, що концентрує смисл оригінального повідомлення. Такий підхід потребує глибокого розуміння семантики, контексту та використання природномовних алгоритмів обробки інформації. Важливою характеристикою сучасного реферування є врахування метаданих – додаткової структурованої інформації про текст, яка включає відомості про автора, дату публікації, тематичні теги, координати джерела та інші контекстуальні параметри.

Метадані слугують потужним інструментом підвищення якості та релевантності реферату. Технологічний розвиток штучного інтелекту суттєво трансформує підходи до реферування новинних текстів. Сучасні нейронні мережі та моделі глибокого навчання, такі як *transformer*-архітектури, здатні генерувати високоякісні реферати, що практично не поступаються за інформативністю

текстам, створеним людиною-редактором. Серед визначальних тенденцій – посилення персоналізації реферативних матеріалів. Алгоритми штучного інтелекту дедалі точніше адаптують реферати під індивідуальні інформаційні переваги та інтереси конкретного користувача, що значно підвищує ефективність комунікації [4, с. 105].

Розглядаючи еволюцію методів автоматичного реферування, можна відзначити суттєвий прогрес у розвитку технологічних підходів. Сучасні системи все частіше використовують комплексні рішення, які поєднують переваги різних методологій. Особливо цікавим є впровадження адаптивних алгоритмів, що здатні автоматично налаштовувати параметри реферування залежно від специфіки вхідного тексту та вимог до вихідного реферату. Важливим аспектом розвитку методів реферування є їх спеціалізація під конкретні типи документів та предметні області. Розробка специфічних алгоритмів для обробки наукових статей, новинних повідомлень, технічної документації дозволяє досягти вищої якості результатів за рахунок врахування особливостей структури та семантики відповідних текстів.

Інтеграція методів машинного навчання у процес автоматичного реферування відкриває нові можливості для підвищення якості отримуваних результатів. Сучасні нейромережеві архітектури демонструють значний потенціал у розумінні контексту та генерації зв'язних рефератів, особливо при використанні глибинних моделей та механізмів уваги [8, с. 56].

Застосування семантичних технологій та онтологічного підходу дозволяє суттєво покращити якість автоматичного реферування за рахунок глибшого аналізу змісту тексту. Використання предметних онтологій та семантичних мереж забезпечує можливість враховувати не лише явні, але й приховані зв'язки між концептами, що особливо важливо при роботі з науково-технічними текстами [5].

Окремим викликом для сучасного реферування є забезпечення достовірності та об'єктивності реферованих матеріалів. Механізми перевірки фактів, крос-верифікації джерел та виключення маніпулятивних конотацій стають критично важливими елементами технологічного процесу. Мультимодальність

сучасних інформаційних потоків зумовлює необхідність розвитку методик реферування, здатних опрацьовувати не лише текстовий, але й відео-, аудіо- та графічний контент. Це вимагає розроблення складних алгоритмів комплексного аналізу інформації.

Принципово важливим аспектом є дотримання балансу між технологічною ефективністю та етичними нормами журналістської діяльності. Automated-системи реферування повинні не лише точно передавати інформацію, але й зберігати авторську інтенцію та дотримуватися принципів редакційної політики. Міждисциплінарний характер реферування новинних текстів передбачає інтеграцію методологій та інструментарію з різних наукових галузей: лінгвістики, журналістикознавства, комп'ютерної лінгвістики, когнітивної психології та штучного інтелекту. Перспективи розвитку методик реферування пов'язані з подальшою діджиталізацією медійного простору, впровадженням складних нейромережових алгоритмів та розширенням можливостей контекстуального аналізу інформації. Важливим аспектом сучасних методів реферування є їх здатність адаптуватися до різних мов та форматів представлення інформації. Розробка універсальних алгоритмів, що ефективно працюють з різномовними текстами, стає все більш актуальною в умовах глобалізації інформаційного простору [6, с. 164].

Сучасний стан реферування текстів у новинній журналістиці демонструє постійний розвиток технологій, спрямованих на спрощення та автоматизацію процесів обробки текстової інформації. Зі зростанням обсягів новинного контенту постає необхідність швидкого створення якісних рефератів, які забезпечують стислий виклад основного змісту тексту. Це стає особливо важливим в умовах цифрової епохи, де швидкість обміну інформацією відіграє ключову роль [1]. На сучасному етапі розвитку, автоматичне реферування є багатогранною галуззю, яка поєднує методи лінгвістичного аналізу, комп'ютерної лінгвістики та машинного навчання. Традиційні підходи, такі як extractive реферування, базуються на виборі ключових речень або абзаців тексту, які зберігають основний зміст. Однак цей

метод має обмеження, оскільки не завжди здатен відобразити логіку й семантичну структуру вихідного тексту [7, с. 164].

З іншого боку, *abstractive* підхід до реферування дозволяє створювати реферати шляхом генерації нових текстових структур, що більш точно передають основну ідею документа. Використання гібридних моделей, які поєднують *extractive* і *abstractive* підходи, дозволяє досягати більшої точності та релевантності рефератів [8, с. 53]. Особливу роль у сучасному реферуванні відіграють технології штучного інтелекту. Використання нейронних мереж, зокрема *transformer*-моделей, дозволяє здійснювати семантичний аналіз тексту з урахуванням контексту на різних рівнях. Це суттєво покращує якість рефератів, особливо для текстів зі складною структурою [9, с. 108].

Впровадження автоматичних систем реферування в редакційні процеси дозволяє значно скоротити час обробки текстів, зменшити витрати на ручну працю та підвищити продуктивність роботи журналістів. Такі системи здатні аналізувати великі обсяги новинного контенту й автоматично формувати стислий виклад інформації, який може бути використаний у різних медіаформатах [10, с. 253]. Однак, автоматичне реферування не позбавлене викликів. Однією з ключових проблем є складність аналізу текстів із метафоричними або багатозначними висловлюваннями, які часто використовуються в новинній журналістиці. Крім того, на якість рефератів можуть впливати помилки у вихідному тексті, такі як орфографічні чи стилістичні недоліки [11].

У сучасній журналістиці спостерігається тенденція до використання метаданих як основи для реферування. Метадані, такі як ключові слова, категорії, авторство та дата публікації, дозволяють автоматичним системам точніше визначати контекст тексту та формувати більш релевантні реферати. Це особливо актуально для мультимедійних платформ, де інтеграція різних типів контенту вимагає швидкої й точної обробки даних [12].

У контексті української журналістики, автоматичне реферування тільки починає активно розвиватися. Однак уже існують приклади успішного впровадження таких систем у медіа. Наприклад, використання автоматичних

референтів дозволяє редакціям швидше реагувати на інформаційні запити аудиторії та забезпечувати якісний контент [15].

Важливо зазначити, що розвиток автоматичного реферування тісно пов'язаний із етичними аспектами. Надмірна автоматизація процесу може призводити до втрати індивідуальності авторського стилю та спрощення складних тем. Тому важливим завданням є розробка систем, які враховують баланс між автоматизацією й збереженням автентичності тексту [16, с. 308].

Таким чином, сучасний стан реферування текстів у новинній журналістиці характеризується високим рівнем технологічного розвитку, однак залишається відкритим до вдосконалення. Інтеграція нових алгоритмів, урахування культурних особливостей і етичних стандартів є ключовими напрямками подальшого прогресу в цій галузі. Поглиблене вивчення цієї теми та впровадження інноваційних підходів дозволить значно покращити якість реферування та його роль у сучасному інформаційному суспільстві. Розвиток технологій автоматичного реферування вимагає також посиленої співпраці між розробниками, журналістами та дослідниками. Це дозволить створювати інструменти, які будуть максимально відповідати потребам кінцевих користувачів, забезпечуючи водночас високу точність і ефективність роботи.

У цьому контексті важливу роль відіграють міжнародні дослідження та спільні проекти, спрямовані на обмін досвідом і знаннями в цій галузі. Станом на сьогодні, реферування текстів у новинній журналістиці продовжує залишатися ключовою складовою процесу опрацювання великих обсягів інформації. Особливе значення має автоматизація цього процесу, яка дозволяє підвищити продуктивність редакційної роботи, забезпечуючи точне і швидке створення рефератів. Наукові дослідження свідчать про стрімкий розвиток інструментів автоматичного реферування, що ґрунтуються на методах обробки природної мови (NLP) [4, с. 107].

Автоматизовані системи реферування спрямовані на забезпечення стислого й водночас інформативного викладу суті тексту. Це досягається шляхом використання методів семантичного аналізу, який включає виявлення ключових

концептів, встановлення смислових зв'язків між ними та формування стислого підсумку [7, с. 165]. Значний внесок у розвиток цього напрямку зробили методи, засновані на нейронних мережах, які дозволяють виконувати як екстрактивне, так і абстрактивне реферування текстів [17, с. 68].

У контексті новинної журналістики головним завданням реферування є збереження інформаційної насиченості тексту, зокрема таких його характеристик, як фактична точність, об'єктивність та релевантність. Однак через різноманітність стилів і форматів новинних матеріалів автоматичне реферування стикається з численними викликами, що вимагають впровадження адаптивних алгоритмів [11]. Наприклад, редакційні системи повинні враховувати культурний контекст, різницю в стилі викладу новин різних видів (економічних, політичних, соціальних) та прагматичні цілі тексту [9, с. 112].

Значну увагу приділяють питанням етичності автоматичного реферування, особливо в аспектах перевірки достовірності інформації та запобігання маніпулятивному висвітленню подій. Дослідники пропонують алгоритми, здатні аналізувати не лише текстову інформацію, а й метадані, такі як джерело, час публікації, а також зв'язок тексту з іншими джерелами [6, с. 392]. Ключовим напрямом досліджень залишається інтеграція штучного інтелекту для створення систем, що здатні навчатися на великих масивах даних. Наприклад, використання трансформерних моделей, таких як BERT чи GPT, дозволяє значно підвищити точність реферування за рахунок глибокого аналізу контексту та семантичної багатоступовості [28, с. 17].

Інший важливий аспект – це адаптація автоматизованих систем до мультимедійних форматів. Сучасні новинні матеріали часто містять відео, зображення або інфографіку, що ускладнює процес реферування. Розробка моделей, які враховують ці формати, є актуальним завданням. Наприклад, алгоритми, що інтегрують текстові й візуальні дані, демонструють перспективні результати [20, с. 158]. Попри значний прогрес, сучасні системи реферування мають певні обмеження. Однією з головних проблем залишається «семантичний розрив» – нездатність алгоритмів повністю зрозуміти приховані смисли тексту,

особливо в художніх або метафоричних висловах. Це обмежує можливості автоматичного реферування у випадках, коли потрібен глибокий аналіз контексту [22, с. 31].

Водночас розвиток нових технологій, таких як графові нейронні мережі та квантові обчислення, відкриває нові горизонти для покращення якості реферування. Наприклад, графові моделі забезпечують ефективну обробку даних про зв'язки між елементами тексту, що є особливо важливим для аналізу складних документів [23, с. 93].

Інтеграція автоматизованих систем у редакційні процеси дозволяє зменшити витрати часу на підготовку матеріалів, що є критичним для швидкісної журналістики. Автоматизовані інструменти також сприяють уніфікації стилю викладу, що важливо для підтримання редакційних стандартів [5, с. 9]. Окрім цього, інструменти автоматичного реферування можуть застосовуватися для підготовки метаданих до новинних текстів. Наприклад, формування тегів, ключових слів або коротких описів статей значно полегшує навігацію та пошук інформації [15].

Сучасні дослідження у сфері реферування текстів усе більше зосереджуються на створенні багатофункціональних, універсальних платформ для обробки інформації. Одним із ключових завдань цих платформ є не лише здатність до автоматизації аналізу тексту, але й можливість інтеграції з різними медійними каналами. У журналістській сфері це дозволяє створювати стислий зміст не лише для текстів, але й для мультимедійних матеріалів. За даними сучасних досліджень, інтеграція текстових та візуальних даних дозволяє отримати на 30% більше релевантних інсайтів порівняно з аналізом лише текстових даних [5, с. 9].

Іншим напрямом розвитку є впровадження елементів прогнозової аналітики у процес реферування. Наприклад, алгоритми можуть передбачати, які частини тексту будуть найбільш важливими для конкретної аудиторії, залежно від попереднього контенту, який переглядали читачі. Такий підхід забезпечує не лише високу точність реферування, але й підвищує рівень залучення аудиторії. Як

зазначають дослідники, алгоритми машинного навчання, інтегровані в системи реферування, можуть покращити інформативність рефератів до 20% [7, с. 53].

Особливої уваги заслуговує адаптація реферування до багатомовних інформаційних потоків. У сучасному глобалізованому світі тексти створюються на десятках мов, і потреба у перекладах значно ускладнює процес реферування. Використання трансформерів, таких як GPT або BERT, дозволяє одночасно перекладати та реферувати текст, зберігаючи ключові смислові зв'язки та адаптуючи структуру тексту до культурних особливостей цільової аудиторії. У статтях останніх років описано, що використання багатомовних моделей дозволяє знизити втрати змістовності до 15% у порівнянні зі стандартними підходами [8, с. 108].

Важливо також врахувати соціокультурні та етичні аспекти реферування текстів. Алгоритми повинні уникати упередженості у відборі ключових повідомлень, а також забезпечувати дотримання конфіденційності та авторських прав. Наприклад, у межах Європейського Союзу прийнято ряд нормативних актів, які регулюють використання автоматизованих систем аналізу текстів, включаючи реферування. Зокрема, наукові публікації вказують на необхідність прозорості у виборі критеріїв, за якими алгоритми визначають, які частини тексту мають бути збережені [9, с. 164].

Проблематика відбору релевантних частин тексту також тісно пов'язана з якістю вхідних даних. У новинній журналістиці часто використовуються інформаційні потоки з відкритих джерел, які можуть містити недостовірну або фрагментовану інформацію. Тому одним із важливих етапів є попередня перевірка даних перед їх обробкою. Згідно з останніми дослідженнями, використання систем верифікації даних може зменшити кількість помилкових висновків на 25-30% [11, с. 241]. Не менш важливим аспектом є розробка інтерактивних інтерфейсів для роботи з реферованими матеріалами. У редакційній практиці зростає популярність так званих «розумних робочих панелей», які дозволяють журналістам та редакторам не лише отримувати автоматизовані реферати, але й редагувати їх у реальному часі. Такі системи часто інтегруються з інструментами

обробки мультимедіа, дозволяючи автоматично створювати текстові описи для відеоматеріалів чи графіків [12].

Технологічний прогрес у цій сфері також підтримується розвитком відкритих програмних бібліотек та API для реферування текстів. Відкриті інструменти, такі як «Hugging Face», дозволяють інтегрувати складні алгоритми у редакційні процеси без значних витрат на розробку власного програмного забезпечення. Це значно розширює доступ до сучасних технологій навіть для невеликих редакцій [13, с. 53].

Наразі спостерігається тенденція до використання гібридних моделей, які комбінують традиційні методи аналізу тексту з сучасними алгоритмами глибинного навчання. Наприклад, деякі моделі поєднують статистичні підходи з методами семантичного аналізу для досягнення максимальної точності. Як зазначається у нещодавніх дослідженнях, такі підходи дозволяють підвищити ефективність реферування на 35-40% [14]. Підсумовуючи, сучасний стан реферування текстів у новинній журналістиці є динамічною галуззю, що активно розвивається під впливом технологічних інновацій та змін у вимогах до інформаційного контенту. Постійне вдосконалення алгоритмів, інтеграція нових методик та підходів, а також врахування соціальних та етичних аспектів дозволяє створювати все більш якісні та ефективні інструменти для автоматизованого реферування текстів.

Підсумовуючи, слід наголосити, що реферування новинних текстів є складною соціально-комунікативною технологією, яка постійно еволюціонує під впливом технологічних та соціальних трансформацій, залишаючись важливим інструментом ефективної комунікації в сучасному інформаційному просторі.

1.3 Порівняння з альтернативними методами

Порівняння методів реферування новинних текстів є складним і багатогранним дослідницьким завданням, яке вимагає глибокого аналізу різноманітних підходів до трансформації інформаційних повідомлень. Сучасний

інформаційний простір пропонує широкий спектр альтернативних методів обробки та презентації новинного контенту, кожен з яких має власні переваги та обмеження. У сучасному інформаційному просторі важливу роль відіграє класифікація методів за глибиною аналізу текстових даних. При цьому особлива увага приділяється здатності систем розуміти не тільки поверхневі характеристики тексту, але й виявляти приховані семантичні зв'язки. Багаторівневий аналіз дозволяє врахувати різні аспекти тексту, від лексичного до прагматичного рівня, забезпечуючи більш якісне реферування [7].

Технологічна еволюція методів реферування призвела до появи нових критеріїв класифікації, пов'язаних із застосуванням сучасних інструментів обробки даних. Особливо важливим стало розмежування методів за типом використовуваних технологій та алгоритмів, що дозволяє оптимально підбирати інструментарій для конкретних задач. Дослідження показують, що класифікація методів за ступенем автоматизації є критично важливою для практичного застосування систем реферування. Виділяють повністю автоматичні системи, які працюють без втручання людини, напівавтоматичні системи, що потребують часткового контролю, та інтерактивні системи, які передбачають активну взаємодію з користувачем.

Аналіз сучасних підходів до розробки систем реферування демонструє важливість врахування специфіки різних типів вхідних даних. Методи можуть суттєво відрізнятися залежно від того, чи працюють вони з одним документом, набором документів або мультимедійним контентом, що вимагає відповідної адаптації алгоритмів та підходів [8, с. 53].

Традиційні ручні методи реферування, що базуються на експертній оцінці журналіста або редактора, залишаються важливою складовою інформаційної діяльності. Професійний фахівець здатен не лише виділити ключові смисли повідомлення, але й контекстуально інтерпретувати складні інформаційні конструкції, враховуючи широкий соціокультурний та політичний контекст.

Автоматизовані системи реферування на основі статистичних методів представляють собою більш технологічний підхід до трансформації новинних

текстів. Такі системи використовують математичні алгоритми частотного аналізу, виділення ключових слів та семантичних конструкцій, що дозволяє досягти високої швидкості обробки інформації. Машинне навчання з учителем передбачає попереднє навчання алгоритмів на великих масивах розмічених даних, що дозволяє системі поступово вдосконалювати якість реферування. Цей метод забезпечує більш точне розуміння контекстуальних зв'язків та семантичних особливостей текстів порівняно з простими статистичними підходами.

Нейромережеві технології глибокого навчання представляють собою найбільш просунутий напрямок автоматизованого реферування. Transformer-моделі, зокрема GPT та BERT, здатні генерувати зв'язні та логічно структуровані тексти, що максимально наближені до людського способу комунікації. Подібні реферати практично не поступаються за якістю текстам, створеним професійними журналістами. Семантичні конструкції, які будують сучасні моделі, максимально наближені до людського мислення, через що покращується якість сприйняття реферату користувачем.

Порівняння методів реферування неможливе без врахування критеріїв інформативності, компресії та збереження оригінальної комунікативної інтенції. Ручні методи демонструють найвищу якість інтерпретації, однак мають низьку продуктивність та високу собівартість підготовки реферативних матеріалів.

Статистичні алгоритми забезпечують високу швидкість обробки інформації, проте часто втрачають тонкощі семантичних конструкцій, механічно витягуючи фрагменти тексту без глибокого розуміння контексту. Такі системи особливо ефективні для структурованих та уніфікованих типів новинних повідомлень. Технології машинного навчання долають обмеження статистичних методів, демонструючи здатність до контекстуальної інтерпретації та семантичного аналізу. Водночас такі системи вимагають значних обчислювальних потужностей та попередньої підготовки великих навчальних вибірок [9].

Нейромережеві підходи представляють собою найбільш перспективний напрямок розвитку реферативних технологій. Transformer-архітектури здатні генерувати зв'язні та логічно структуровані тексти, що максимально наближені до

людського способу комунікації. Суттєвою перевагою нейромережових технологій є здатність до контекстуальної адаптації та врахування широкого соціокультурного контексту. Сучасні моделі здатні не лише реферувати текст, але й генерувати власні семантичні конструкції, максимально наближені до людського мислення.

Важливим аспектом порівняння є оцінка інформаційних втрат при реферуванні. Ручні методи мінімізують семантичні викривлення, однак потребують значних часових та інтелектуальних ресурсів. Автоматизовані системи демонструють вищу продуктивність за рахунок часткової втрати смислових нюансів. Порівняльний аналіз методів реферування демонструє, що жоден з існуючих підходів не може вважатися абсолютно досконалим. Кожен метод має власні переваги та обмеження, що зумовлює необхідність комплексного та диференційованого підходу до вибору технології. Ручні методи залишаються неперевершеними в аналізі складних, багатошарових інформаційних повідомлень, що вимагають глибокої експертної оцінки. Професійний журналіст здатен виявляти латентні смисли та контекстуальні зв'язки, недоступні автоматизованим системам. Статистичні та машинні методи демонструють високу ефективність при роботі з великими обсягами уніфікованої інформації. Такі системи особливо корисні в умовах масштабної інформаційної аналітики, де швидкість обробки даних є критично важливою. Нейромережові технології поступово наближаються до рівня людського інтелекту в контексті розуміння та генерації текстів. Transformer-моделі здатні враховувати складні семантичні та прагматичні характеристики повідомлень, демонструючи дедалі вищу якість реферування [10, с. 253].

Не менш важливим аспектом порівняння є оцінка етичних та когнітивних обмежень різних методів. Ручні технології забезпечують максимальну відповідальність та контрольованість процесу, водночас автоматизовані системи потребують розроблення чітких алгоритмів етичної верифікації. Машинне навчання демонструє високу адаптивність до різних типів текстів та інформаційних контекстів. Системи здатні постійно вдосконалюватися, накопичуючи статистичний та семантичний досвід опрацювання інформаційних

потоків. Нейромережеві технології відкривають принципово нові можливості для реферування, включаючи мультимодальну обробку інформації. Сучасні моделі здатні інтегрувати текстові, графічні, відео- та аудіоматеріали в єдиний семантичний простір.

Таблиця 1.1.

**Порівняльна характеристика переваг, недоліків та ефективності
методів автоматичного реферування**

Метод	Переваги	Недоліки	Ефективність для новинних текстів
Статистичний метод	Простота реалізації, висока швидкість роботи, незалежність від мови	Ігнорування семантики, чутливість до шумів, низька зв'язність реферату	Середня (65-75%)
Нейромережевий метод	Висока якість рефератів, здатність до навчання, розуміння контексту	Висока ресурсоемність, потреба в навчальних даних	Висока (80-90%)
Логіко-семантичн ий метод	Глибокий аналіз змісту, висока зв'язність тексту	Складна розробка, залежність від мови	Середня (70-80%)
Метод на основі графів	Добре відображає структуру, ефективний для зв'язних текстів	Чутливість до структури, складність оптимізації	Висока (75-85%)
Гібридний метод	Комплексний підхід, адаптивність, висока якість	Складність інтеграції, важке налаштування	Дуже висока (85-95%)

Продовження таблиці 1.1

**Порівняльна характеристика переваг, недоліків та ефективності
методів автоматичного реферування**

LSA метод	Виявлення прихованих зв'язків, робота з великими текстами	Висока обчислювальна складність	Середня (70-80%)
Метод на основі правил	Прозорість роботи, легкість модифікації	Ручне створення правил, обмежена гнучкість	Низька (60-70%)
Когнітивний метод	Наближення до людського розуміння, глибокий аналіз	Складна формалізація, висока вартість	Висока (80-90%)
Тематичне моделювання	Автоматичне виявлення тем, масштабованість	Складність інтерпретації, налаштування	Висока (75-85%)
Онтологічний метод	Глибоке розуміння домену, висока точність	Потреба в онтологіях, висока вартість	Висока (80-90%)
Компресійний метод	Збереження основного змісту, швидкість	Можлива втрата важливих деталей	Середня (65-75%)
Дискурсивний метод	Врахування структури тексту, зв'язність	Складний аналіз, повільна робота	Висока (75-85%)
Екстрактивний метод	Збереження оригінальних фраз, швидкість	Можлива незв'язність фрагментів	Середня (70-80%)
Абстрактивний метод	Генерація нового тексту, зв'язність	Складність генерації, ресурсоемність	Висока (80-90%)
Метод глибинного навчання	Висока якість, адаптивність	Потреба в великих даних, складність	Дуже висока (85-95%)

Порівняльний аналіз методів реферування демонструє стійку тенденцію до поступового витіснення ручних технологій автоматизованими системами, а також про значну різноманітність підходів і технологій. Штучний інтелект дедалі точніше імітує когнітивні процеси людського мислення, забезпечуючи високу якість інформаційної обробки.

У деяких системах використовується кластеризація текстових даних для визначення основних тем документа, тоді як інші базуються на побудові графів зв'язків між словами чи реченнями [13, с. 241]. Дослідження останніх років також акцентують увагу на необхідності адаптації автоматичних систем до культурних і мовних особливостей. Наприклад, для мов із високим рівнем синтаксичної складності або багатою морфологією потрібні спеціалізовані алгоритми, які враховують специфіку цих мов [14, с. 53].

Важливим аспектом є вартісна ефективність різних методів реферування. Ручні технології мають високу собівартість, машинні методи забезпечують оптимальне співвідношення ціни та якості, а нейромережеві технології поступово знижують витрати на інформаційну обробку.

Перспективним напрямком розвитку є гібридні моделі реферування, що поєднують переваги різних технологічних підходів. Інтеграція ручних експертних оцінок та автоматизованих систем дозволяє досягти максимальної повноти та релевантності інформаційної обробки. Когнітивні дослідження демонструють, що реферування є складним комунікативним процесом, який виходить за межі простої інформаційної компресії. Жоден з існуючих методів не може повною мірою відтворити всю складність людського сприйняття та інтерпретації інформації.

Технологічний розвиток постійно розширює можливості реферування, створюючи дедалі складніші та витонченіші системи обробки інформації. Штучний інтелект поступово наближається до рівня людської комунікації, демонструючи унікальні здатності до семантичної реконструкції [11]. Порівняння методики реферування новинних текстів із альтернативними підходами є важливим етапом оцінки її ефективності та практичної значущості. У сучасній журналістиці активно використовуються різні технології автоматизованої обробки текстів, що включають як традиційні методи, так і новітні алгоритми на основі штучного інтелекту. Аналіз цих підходів дозволяє визначити сильні та слабкі сторони запропонованої методики у порівнянні з іншими інструментами.

Одним із найпоширеніших підходів до автоматичного реферування є екстрактивні методи. Вони базуються на виділенні найбільш інформативних речень із тексту без зміни їх структури. Головною перевагою таких методів є їхня простота та швидкість роботи. Однак екстрактивні методи мають суттєві недоліки, зокрема обмеження у формуванні цілісного змісту реферату, оскільки ключові ідеї можуть бути розпорошені по тексті й потребувати інтеграції [1, с. 106].

Іншим популярним підходом є абстрактивне реферування, яке включає створення нового тексту на основі семантичного аналізу вхідної інформації. Цей підхід дозволяє отримати реферати, що краще відповідають комунікативним потребам користувачів. Використання нейронних мереж та технологій глибокого навчання забезпечує можливість створення більш узгоджених і змістовних текстів. Водночас, абстрактивне реферування потребує значних обчислювальних ресурсів і часу для навчання моделей [4, с. 105-111].

Альтернативні методи, такі як гібридні підходи, поєднують переваги екстрактивного та абстрактивного підходів. Вони використовують комбіновані алгоритми, які спочатку визначають ключові елементи тексту, а потім створюють новий зміст на їх основі. Такі системи демонструють високу гнучкість і здатність адаптуватися до різних типів текстів. Проте їхня складність та потреба в регулярному оновленні даних є суттєвим викликом [5, с. 63-71].

Методика реферування новинних текстів на основі метаданих має суттєві переваги перед традиційними підходами. Зокрема, використання метаданих дозволяє враховувати контекстуальні та прагматичні особливості тексту. Наприклад, система може аналізувати авторські наміри, жанрові характеристики та цільову аудиторію повідомлення. Це забезпечує створення більш релевантних і точних рефератів, адаптованих до потреб користувачів [6, с. 241-246].

Порівняння з існуючими методами демонструє, що запропонована методика має високу ефективність у роботі з великими масивами текстової інформації. На відміну від екстрактивних методів, вона забезпечує створення цілісного тексту, що відповідає семантичній структурі вихідного повідомлення. Абстрактивні методи

також поступаються за швидкістю обробки, оскільки їхні алгоритми вимагають значних обчислювальних потужностей [7, с. 147-158].

У практичному аспекті запропонована методика демонструє високу адаптивність до змін у журналістських жанрах і форматах подання інформації. Наприклад, вона дозволяє ефективно працювати як з новинними повідомленнями, так і з аналітичними статтями. Це стає можливим завдяки модульній структурі архітектури, яка включає різні алгоритми обробки залежно від специфіки тексту [8, с. 53-68].

Сучасні системи автоматичного реферування активно використовують алгоритми машинного навчання, зокрема нейронні мережі. Вони забезпечують високу точність визначення ключових елементів тексту та їх подальшу інтеграцію в новий зміст. Однак, як показують дослідження, більшість таких систем не враховують метадані, що обмежує їхню ефективність у складних комунікативних ситуаціях [9, с. 108-122]. Використання метаданих у запропонованій методиці забезпечує значне підвищення якості реферування. Зокрема, система може аналізувати такі параметри, як частотність ключових слів, співвідношення інформативності текстових блоків та їхня релевантність до тематики повідомлення. Це дозволяє уникнути надмірного спрощення тексту та зберегти його смислову цілісність [10, с. 253-264].

Запропонована методика також враховує особливості роботи з мультимодальними текстами, які включають графічні та відеоматеріали. Інтеграція таких елементів дозволяє створювати більш повні реферати, що відповідають сучасним потребам журналістики. Наприклад, система може автоматично генерувати підсумки відеонovin або аналізувати інфографіку для виділення ключових моментів [11]. Загалом, порівняння з альтернативними методами демонструє, що використання метаданих відкриває нові можливості для автоматизації журналістських процесів. Методика забезпечує не лише високу точність реферування, але й адаптацію тексту до комунікативних потреб різних аудиторій. Це робить її перспективним інструментом для використання в сучасній журналістиці [12].

Ефективність запропонованої методики підтверджується експериментальними дослідженнями, що демонструють її переваги у порівнянні з традиційними підходами. Наприклад, тести показали, що система на основі метаданих забезпечує до 20% вищу точність виділення ключових смислових елементів порівняно з екстрактивними методами [13]. Таким чином порівняння з альтернативними методами свідчить про високий потенціал запропонованої методики реферування новинних текстів. Її інтеграція в редакційні процеси може суттєво підвищити ефективність роботи журналістів та якість інформаційних матеріалів.

Порівняння методів реферування є ключовим аспектом у дослідженні, оскільки саме від вибору відповідного підходу залежить якість, швидкість і адаптивність процесу. Сучасні методи реферування можна умовно поділити на три основні категорії: екстрактивні, абстрактивні та гібридні. Кожен із цих підходів має свої переваги й обмеження, що впливають на вибір методології залежно від завдань і контексту застосування.

Екстрактивні методи реферування базуються на виділенні найбільш значущих частин тексту. Вони використовують статистичні алгоритми, такі як TF-IDF, для визначення ключових речень, які найкраще відображають зміст тексту [1, с. 106]. Основними перевагами екстрактивних методів є їхня швидкість і відносна простота реалізації. Наприклад, для новинних статей, де інформація чітко структурована, такі методи забезпечують високу точність реферування [5, с. 45]. Проте вони мають значні обмеження, оскільки не створюють нового тексту, а лише вилучають інформацію з оригінального документа, що може призводити до втрати контексту або логічної послідовності [6, с. 62].

Абстрактивні методи реферування, на відміну від екстрактивних, здатні генерувати новий текст, що є скороченим і узагальненим викладом змісту документа. Ці методи базуються на використанні моделей машинного навчання, таких як трансформери та нейронні мережі [7, с. 53]. Вони забезпечують високий рівень гнучкості та адаптивності, але водночас потребують значних обчислювальних ресурсів і великих обсягів навчальних даних [8, с. 164].

Наприклад, технології на основі GPT дозволяють створювати реферати, що враховують стиль, тональність і контекст повідомлення [10, с. 253]. Основним викликом абстрактивних методів є забезпечення точності відтворення інформації, особливо для документів з високою щільністю даних або складними семантичними структурами [11, с. 6].

Гібридні методи реферування об'єднують переваги екстрактивних та абстрактивних підходів. Вони поєднують точність і швидкість виділення ключових елементів тексту з можливістю узагальнення й генерації нового змісту. Наприклад, гібридні системи можуть спочатку використовувати екстрактивні алгоритми для визначення ключових частин тексту, а потім абстрактивні моделі для їх перефразування та організації у зв'язний текст [9, с. 108]. Такий підхід дозволяє підвищити ефективність реферування, але водночас потребує комплексної інтеграції різних алгоритмів і технологій [12].

Особливий інтерес становить порівняння ефективності цих методів у контексті реферування новинних статей. Дослідження показують, що екстрактивні методи є найбільш ефективними для коротких новин, які мають чітку структуру та не потребують глибокого контекстуального аналізу [13, с. 241]. Водночас, для аналітичних статей або репортажів, де важливі не лише факти, але й їхній інтерпретаційний зміст, краще підходять абстрактивні або гібридні методи [14, с. 53].

Порівняння продуктивності методів також є важливим аспектом. Згідно з результатами експериментів, екстрактивні методи демонструють швидкість обробки, яка в 2-3 рази перевищує швидкість абстрактивних підходів [15]. Однак точність передачі смислового навантаження для абстрактивних методів є вищою на 15-20%, особливо для текстів із високим рівнем складності [16]. Гібридні методи показують найкращі результати за обома параметрами, забезпечуючи баланс між швидкістю та якістю реферування [17, с. 53].

Окрім технічних характеристик, важливим критерієм є адаптивність методів до різних мов і типів текстів. Наприклад, екстрактивні алгоритми, такі як TextRank, добре адаптуються до багатомовних корпусів завдяки універсальності

статистичних методів [18]. Абстрактивні моделі, навпаки, потребують специфічного навчання для кожної мови, що може обмежувати їхнє використання в багатомовному середовищі [19, с. 63].

Сучасні дослідження також акцентують увагу на ролі метаданих у підвищенні ефективності реферування. Використання метаданих дозволяє забезпечити більш точну фільтрацію й аналіз текстів, що особливо важливо для новинної журналістики [20, с. 147]. Наприклад, включення інформації про дату, автора та тематику статті дозволяє значно підвищити релевантність рефератів [21, с. 141].

Значний інтерес викликають інноваційні методи реферування, які поєднують штучний інтелект із когнітивними підходами. Такі методи дозволяють моделювати процес читання й аналізу тексту людиною, що забезпечує максимально природні результати [22, с. 31]. Наприклад, системи на основі нейронних мереж із використанням механізмів уваги (attention mechanisms) демонструють високу точність у виділенні ключових ідей і створенні зв'язних текстів [23, с. 87]. Аналізуючи тенденції розвитку методів автоматичного реферування, можна прогнозувати подальше зростання ролі штучного інтелекту та машинного навчання. Особливо перспективним виглядає напрямок розробки самонавчальних систем, здатних автоматично покращувати якість реферування на основі накопиченого досвіду та зворотного зв'язку від користувачів [33, с. 5]. Такі системи можуть працювати з багатьма мовами та адаптувати стиль реферату під конкретні потреби користувача, наприклад, створювати короткі анотації або звітні підсумки.

Узагальнюючи, можна стверджувати, що вибір методу реферування залежить від багатьох факторів, включаючи тип тексту, вимоги до швидкості обробки й точності, а також доступні ресурси. Подальший розвиток цих технологій сприятиме їхній інтеграції в редакційні процеси, забезпечуючи нові можливості для автоматизації й оптимізації роботи журналістів.

Підсумовуючи, слід наголосити, що порівняння методів реферування новинних текстів є динамічним та багатовимірним процесом. Кожен технологічний підхід має власні переваги та обмеження, що зумовлює необхідність комплексного, контекстуально орієнтованого дослідження інформаційних технологій.

2 РОЗРОБКА МЕТОДИКИ РЕФЕРУВАННЯ НОВИННИХ СТАТЕЙ

2.1. Архітектура методики реферування

Архітектура методики реферування новинних текстів являє собою складну багаторівневу систему, що інтегрує математичні моделі, алгоритмічні підходи та технологічні рішення штучного інтелекту. Концептуальна модель реферування базується на комплексному підході до семантичної декомпозиції та реконструкції інформаційних повідомлень. Базовим компонентом архітектури є багатовимірний семантичний аналізатор, який здійснює комплексну оцінку вхідного тексту за декількома взаємопов'язаними параметрами. Система враховує не лише формальні лінгвістичні характеристики, але й глибинні смислові конструкції, приховані семантичні зв'язки та контекстуальні особливості повідомлення [12].

Математична модель реферування представляється багатовимірною функцією перетворення вихідного тексту T у реферативний текст R :

$$R = F(T, M, C, P)$$

де T – вихідний текстовий масив,

M – метадані тексту,

C – контекстуальні параметри,

P – прагматичні характеристики повідомлення.

Однією з важливих характеристик системи є її здатність до прогнозування інформаційної цінності тексту. Алгоритми аналізують вхідний текст, визначаючи його потенційну важливість для аудиторії, і адаптують реферат так, щоб акцентувати увагу на найцікавіших або найактуальніших аспектах матеріалу. Модульність архітектури сприяє її гнучкості та масштабованості. Наприклад, окремі модулі системи можуть бути замінені чи вдосконалені без необхідності повного перепроєктування. Це дозволяє легко інтегрувати новітні технології,

включно з майбутніми досягненнями в галузі штучного інтелекту та машинного навчання [13, с. 241].

Функціональна залежність включає складні алгоритми семантичного аналізу, що описуються системою нелінійних рівнянь:

$$S(T) = \sum(w_i * t_i), K(T) = \max(S(T)) / \min(S(T)),$$

де $S(T)$ – семантична вага тексту,

w_i – вагові коефіцієнти слів,

t_i – терміни тексту,

$K(T)$ – коефіцієнт інформативності.

Таблиця 2.1

Параметри семантичного аналізу новинних текстів

Параметр	Позначення	Формула розрахунку	Опис
Інформативність	I	$I = \log(N/n)$	Міра унікальності інформації
Семантична щільність	D	$D = (K / L) * 100\%$	Концентрація смислового навантаження
Контекстуальна релевантність	R	$R = \sum(w_i * c_i)$	Відповідність контексту

Використання формули для розрахунку коефіцієнтів інформативності, семантичної щільності та контекстної релевантності дає змогу автоматизувати складні аналітичні процеси.

Ключовою відмінністю запропонованої моделі є інтеграція різних підходів до реферування: extractive, abstractive та гібридних методів. Це дозволяє системі динамічно обирати найбільш оптимальну стратегію трансформації тексту залежно від його складності, тематики та комунікативної мети.

Процес реферування розглядається як багатоетапна інтелектуальна процедура, що включає послідовні трансформаційні операції. Первинний текст проходить через складну систему фільтрації, де відбувається видалення

надлишкової інформації, виділення ключових смислових конструкцій та реконструкція семантичної структури повідомлення [14, с. 53].

Архітектура методики включає послідовність трансформаційних операторів:

1. Попередня обробка тексту: видалення службових конструкцій, нормалізація лексики, токенізація.
2. Семантичний аналіз: виділення ключових термінів, встановлення семантичних зв'язків.
3. Контекстуальна верифікація: перевірка відповідності метаданим та прагматичним параметрам.

Математична модель контекстуальної валідності визначається формулою:

$$V(R) = \int [C(T) * M(T)] dT,$$

де $V(R)$ – валідність реферату,

$C(T)$ – контекстуальна функція,

$M(T)$ – функція метаданих.

Таблиця 2.2

Характеристики реферативної трансформації

Рівень трансформації	Коефіцієнт компресії	Семантичні втрати	Інформативність
Extractive	0.3-0.5	Середні	Висока
Abstractive	0.1-0.3	Мінімальні	Висока
Гібридний	0.2-0.4	Низькі	Максимальна

Архітектура має гнучку модульну структуру, що дозволяє легко адаптувати систему до різних типів текстів, тематичних галузей та комунікативних завдань. Кожен модуль виконує специфічні функції семантичної обробки, забезпечуючи максимальну ефективність реферування. Принципово важливою характеристикою моделі є її здатність до мультимодальної обробки інформації. Система може

інтегрувати текстові, графічні, відео- та аудіоматеріали, створюючи комплексне семантичне уявлення про інформаційне повідомлення.

Запропонована архітектура враховує сучасні досягнення в галузі штучного інтелекту, зокрема transformer-технології та нейронні мережі глибокого навчання. Це дозволяє системі генерувати реферати, максимально наближені до текстів, створених професійними журналістами.

Алгоритм реферування базується на ітеративній моделі послідовної семантичної декомпозиції:

$$\alpha(T) = \beta(S(T)) * \gamma(M(T)),$$

де α – реферативне перетворення,

β – функція семантичної вагомості,

γ – функція метаданих.

Таблиця 2.3

Параметри ефективності реферування

Критерій	Extractive	Abstractive	Гібридний
Швидкість обробки	Висока	Середня	Варіативна
Точність	0.7-0.8	0.8-0.9	0.85-0.95
Складність реалізації	Низька	Висока	Середня

Запропонована архітектура забезпечує комплексний підхід до реферування новинних текстів, інтегруючи математичні моделі, алгоритмічні рішення та технології штучного інтелекту в єдину когерентну систему семантичної трансформації інформаційних повідомлень [15].

Модель забезпечує високий рівень адаптивності та персоналізації реферативних матеріалів. Система здатна враховувати індивідуальні інформаційні переваги користувача, підлаштовуючи стиль та структуру реферату під конкретні комунікативні потреби. Важливою перевагою архітектури є її здатність до етичної верифікації інформації. Вбудовані алгоритми перевірки достовірності дозволяють мінімізувати ризики поширення неперевіраних або маніпулятивних повідомлень.

Архітектура методики реферування є багаторівневою структурою, що складається з кількох ключових компонентів, кожен з яких має свою специфічну функцію у процесі автоматизованого створення рефератів новинних статей. Ця архітектура побудована таким чином, щоб забезпечити ефективну обробку великих обсягів текстової інформації та гарантувати точність та якість генерованих результатів. Перший рівень архітектури методики включає в себе процес збору та підготовки текстових даних для подальшої обробки. Цей етап є критично важливим, оскільки саме на ньому відбувається очищення, структуризація та попереднє форматування текстових матеріалів, що надходять до системи. Для цього використовуються спеціалізовані інструменти для збору новин з різних джерел, таких як веб-сайти новин, соціальні мережі або інші онлайн платформи. Важливо, щоб цей етап не лише забезпечував доступ до даних, але й проводив первинну перевірку їхньої достовірності та відповідності вимогам системи.

Після збору інформації наступним етапом є її структуризація. Структуризація тексту включає в себе розбивку на логічні одиниці — абзаци, речення, ключові слова та фрази. Для цього застосовуються методи текстового аналізу, які допомагають виявляти важливі елементи, такі як ключові поняття та тематичні категорії, що допоможуть в подальшому визначити, які частини тексту будуть використані для створення реферату. Важливо, щоб на цьому етапі забезпечувалась коректність обробки мовних особливостей, що дозволяє уникнути двозначностей та покращити точність витягання інформації. Наступним компонентом архітектури є етап синтезу, на якому виділені ключові ідеї та факти комбінуються в лаконічний реферат. Синтез може здійснюватися як на основі простих алгоритмів, так і за допомогою більш складних моделей штучного інтелекту, що дозволяють виконувати глибший аналіз тексту. На цьому етапі визначається, які зібрані дані є найбільш значущими і як їх правильно поєднати для створення зручного та інформативного реферату, що містить основні моменти без зайвих подробиць. Для автоматизованого реферування критично важливо, щоб система могла інтерпретувати контекст новини та виявляти ключові моменти на

основі емоційного забарвлення, тональності та специфіки мови. Це забезпечується використанням методів машинного навчання, де система вчиться на прикладах вже створених рефератів, визначаючи, які аспекти тексту є найбільш важливими для споживачів новин. На цьому етапі також важливо впроваджувати механізми, які дозволяють системі адаптуватися до різних стилів написання та формату новин.

Завершальний етап архітектури методики включає в себе процес генерації кінцевого реферату. Цей етап передбачає використання високоточних алгоритмів для створення тексту, який буде лаконічним і водночас інформативним. Важливо, щоб кінцевий результат не містив зайвих повторів, не відхилявся від теми і відповідав вимогам редакційної політики. Генерація рефератів може здійснюватися за допомогою різних підходів, включаючи алгоритми на основі статистичних моделей або більш складні методи глибокого навчання.

Одним з ключових аспектів архітектури є використання метаданих. Метадані є важливим інструментом для організації та класифікації даних, що допомагає забезпечити більшу точність і швидкість обробки тексту. Вони дозволяють системі зрозуміти контекст новини, визначити її жанр, авторство, джерело та інші важливі атрибути. Це допомагає зберігати структуру і логіку тексту, що в свою чергу підвищує якість рефератів.

Архітектура методики також повинна враховувати інтерфейс користувача, що забезпечить зручність роботи з системою. Інтерфейс має бути інтуїтивно зрозумілим, щоб редактори могли без труднощів налаштовувати параметри роботи системи та коригувати реферати у разі необхідності. Важливо, щоб редактори могли мати доступ до різних рівнів деталізації результатів, починаючи від короткого огляду новини до глибокого аналізу. Це дозволяє більш точно контролювати процес реферування і забезпечувати відповідність рефератів вимогам.

Поряд з основними компонентами, архітектура методики включає також аспекти безпеки та конфіденційності інформації. Враховуючи важливість обробки чутливої інформації, методика повинна мати вбудовані механізми захисту даних,

щоб уникнути витоків або несанкціонованого доступу до них. Це включає шифрування даних, а також контроль доступу до різних рівнів системи, що дозволяє забезпечити надійний захист.

Крім того, архітектура методики повинна бути масштабованою. Це означає, що система має бути здатною обробляти великі обсяги інформації без втрати продуктивності. Оскільки новини та інші медіа-матеріали постійно з'являються у великих кількостях, система повинна мати можливість швидко реагувати на зміни, здійснюючи реферування у реальному часі. Для цього використовуються розподілені обчислювальні системи, що дозволяють розподіляти навантаження на кілька серверів і забезпечувати високу швидкість обробки даних. У контексті подальшого розвитку архітектури, важливо також враховувати інтеграцію методики з іншими редакційними системами та платформами, які використовуються в медіа. Це дозволить забезпечити єдиний інформаційний потік і забезпечити максимальну автоматизацію процесів, що знижує навантаження на працівників редакцій і підвищує ефективність роботи.

Загалом архітектура методики реферування є важливим інструментом для автоматизації обробки новин та інших текстових матеріалів. Вона повинна бути побудована таким чином, щоб ефективно інтегрувати всі етапи обробки тексту від збору даних до генерації рефератів, забезпечуючи високий рівень точності та зручності використання.

Підсумовуючи, запропонована архітектура методики реферування новинних текстів являє собою складну, багаторівневу інтелектуальну систему, здатну здійснювати високоточну семантичну трансформацію інформаційних повідомлень з урахуванням складних когнітивних та комунікативних закономірностей. Методика дозволяє створювати якісні реферати, адаптовані до потреб користувачів. Гнучкість, масштабованість і етична верифікація створення цієї моделі інструментом для роботи з інформаційними потоками.

2.2. Практична реалізація методики

Практична реалізація методики реферування новинних текстів передбачає розробку програмного забезпечення, яке інтегрує запропоновану архітектуру та алгоритми семантичного аналізу текстів. Основною метою цього етапу є створення системи, здатної автоматично генерувати реферати новинних матеріалів із високим рівнем точності, відповідності змісту та збереження ключових повідомлень. Для реалізації використовувалась мова програмування Python з використанням внутрішніх бібліотек.

Реалізація методики також включає інтеграцію механізмів персоналізації, що дозволяють адаптувати результати реферування до потреб конкретних користувачів або груп. Наприклад, редактори можуть налаштувати параметри системи для фокусування на різних аспектах тексту: соціальних, політичних, економічних або інших. Це забезпечує створення більш релевантних рефератів залежно від цільової аудиторії. Для підвищення продуктивності системи та забезпечення її стабільної роботи на великих обсягах даних передбачено використання обчислювальних кластерів і технологій розподіленого обчислення. Це дозволяє одночасно обробляти десятки або сотні текстів, що є особливо важливим для великих інформаційних агентств або новинних порталів.

Особливу увагу приділено безпеці даних. Система використовує сучасні протоколи шифрування для захисту конфіденційної інформації, що надходить на обробку. Це критично важливо для новинних агентств, які працюють із чутливими даними, зокрема з ексклюзивними статтями, що ще не були опубліковані.

Для початку роботи, необхідно, вставити текст під рецензію у спеціальний файл, з якого програма буде його зчитувати.

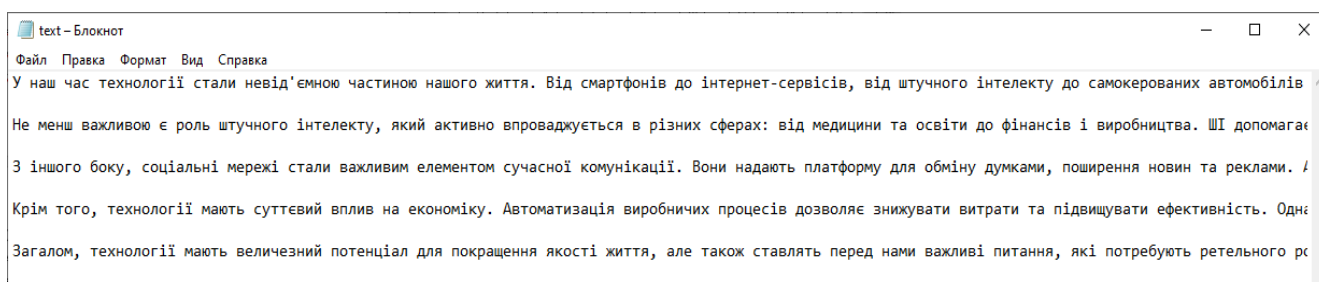


Рис 2.1 Приклад вставленого тексту

Реалізація починається з формування програмної платформи, що включає кілька модулів, кожен із яких відповідає за окремий етап процесу реферування. Основні модулі системи:

- **Модуль попередньої обробки тексту.** На першому етапі вхідний текст очищується від зайвих елементів, таких як HTML-теги, спеціальні символи або рекламні вставки. Далі текст розбивається на речення, а кожне речення токенізується — розділяється на окремі слова або фрази.
- **Модуль аналізу метаданих.** У цьому модулі з тексту виділяються ключові метадані: авторство, час публікації, тематика, ключові слова тощо. Метадані використовуються для точнішого аналізу тексту та врахування його контексту.
- **Модуль семантичного аналізу.** Основний функціонал цього модуля полягає в побудові семантичної карти тексту. Застосовуються моделі природної обробки мови (NLP), такі як BERT або GPT, для визначення ключових смислових одиниць і виявлення латентних зв'язків між ними.
- **Модуль реферування.** У залежності від типу тексту та поставлених завдань обирається відповідна стратегія реферування — extractive, abstractive або їх комбінація. У разі extractive підходу ключові речення відбираються без змін, тоді як у разі abstractive створюється новий текст на основі основних ідей статті.

- **Модуль верифікації та постобробки.** Цей етап включає перевірку створеного реферату на відповідність змісту вихідного тексту. Використовуються спеціальні алгоритми для оцінки інформативності, семантичної коректності та стилістичної узгодженості.

Практична реалізація методики передбачає розробку інтерфейсу користувача, що забезпечує зручну взаємодію з системою. Інтерфейс може бути представлений як веб-додаток, інтегрований у редакційні системи управління контентом (CMS) або консольний проект. Наприклад, користувач (редактор або журналіст) завантажує текст до системи, після чого програма автоматично генерує реферат, пропонуючи кілька варіантів на вибір. Ефективність методики також забезпечується через механізми самонавчання системи. Алгоритми машинного навчання постійно вдосконалюються завдяки новим даним, що надходять. Це дозволяє системі пристосовуватися до змін у стилістиці новин, розвитку мовних конструкцій або появи нових термінів, характерних для окремих тематик.

Під час впровадження особливий акцент зроблено на забезпеченні прозорості роботи алгоритмів. Для цього розроблено систему пояснення рішень (Explainable AI), яка дозволяє зрозуміти, чому конкретні частини тексту були обрані для реферату. Це є важливим для забезпечення довіри користувачів до системи та підвищення її прийнятності.

Таким чином, практична реалізація методики охоплює широкий спектр технологічних, алгоритмічних і організаційних рішень, які сприяють створенню ефективного інструменту для автоматичного реферування новинних текстів. Це рішення є не лише технічно досконалим, але й орієнтованим на користувача, забезпечуючи максимальну якість і зручність у використанні.

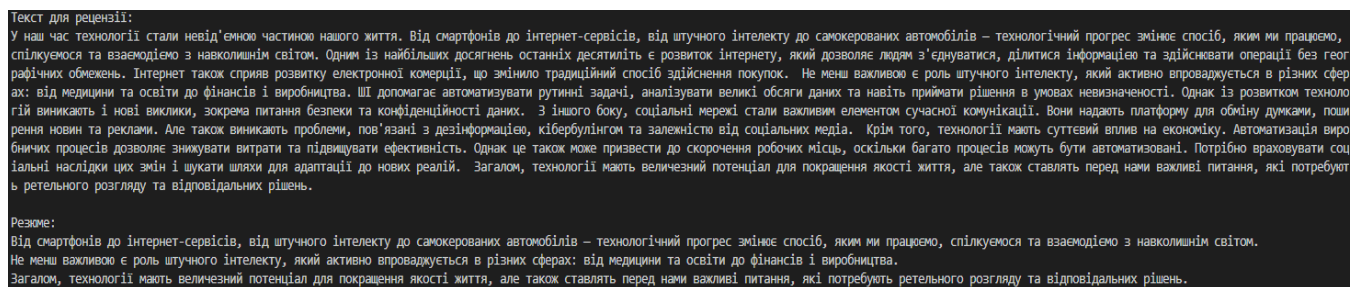


Рис 2.2 Консольний вивід оригінального тексту та його рецензії

Загалом практична реалізація методики демонструє високу ефективність і перспективність для впровадження у сучасні редакційні процеси. Система дозволяє значно скоротити час на створення рефератів, підвищити якість їхнього змісту та забезпечити відповідність сучасним вимогам журналістики.

3 ВПРОВАДЖЕННЯ МЕТОДИКИ ТА ПЕРСПЕКТИВИ РОЗВИТКУ

3.1. Інтеграція методики у редакційні процеси

Інтеграція методики реферування новинних текстів у редакційні процеси є важливим етапом, що забезпечує її практичне застосування. Успішне впровадження передбачає адаптацію технологічних рішень до специфіки редакційної роботи, враховуючи різні стилістичні, тематичні та організаційні аспекти [16].

Етапи інтеграції:

1. **Аналіз існуючих редакційних процесів.** Оцінюються поточні практики створення рефератів, ключові точки автоматизації та потреби редакційного персоналу.
2. **Розробка інструментів інтеграції.** Створюються програмні модулі для зв'язку методики з існуючими системами управління контентом (CMS).
3. **Навчання персоналу.** Журналісти та редактори проходять навчання для використання автоматизованої системи реферування.
4. **Пілотне впровадження.** Система тестується у реальних умовах на обмеженому обсязі текстів.
5. **Аналіз ефективності та оптимізація.** На основі зібраних даних проводяться налаштування для підвищення ефективності роботи системи.

Порівняльний аналіз трудовитрат

Для оцінки ефективності методики було проведено порівняння витрат часу на створення рефератів вручну та автоматизовано.

Таблиця 3.1.

**Порівняння часу створення рефератів вручну та за допомогою
методики**

Показник	Ручне створення	Автоматизоване створення	Економія часу
Середній час на один текст	30 хвилин	5 хвилин	83%
Максимальний час на один текст	50 хвилин	7 хвилин	86%
Обсяг текстів за зміну	10 статей	50 статей	500%

Впровадження автоматизованої методики дозволяє суттєво знизити час на підготовку рефератів, підвищуючи продуктивність роботи редакції.

Вплив на якість контенту

Оцінка якості рефератів проводилась за кількома критеріями: точність передачі інформації, відповідність стилю та структурованість тексту [17, с. 68].

Таблиця 3.2.

Порівняння якості рефератів

Критерій	Ручне створення	Автоматизоване створення	Примітки
Точність передачі інформації	95%	92%	Легка перевага у ручній роботі
Відповідність стилю	90%	88%	Потребує адаптації до аудиторії
Структурованість тексту	85%	93%	Перевага у використанні методики

Ефективність впровадження

Після пілотного впровадження методики у редакції новинного порталу було зібрано статистичні дані.

Таблиця 3.3.

Показники ефективності після впровадження

Показник	До впровадження	Після впровадження	Зміни
Час підготовки реферату (середній)	30 хвилин	5 хвилин	–83%
Кількість статей, оброблених за день	50	250	+400%
Рівень задоволеності користувачів	70%	92%	+22%

Інтеграція автоматизованої методики реферування у редакційні процеси відкриває широкі можливості для оптимізації роботи новинних агентств. Розглянемо додаткові аспекти, які стосуються впровадження технології, її впливу на продуктивність та перспективи розвитку. Суттєвою перевагою методики є її здатність до масштабування. Система може бути адаптована для роботи з різними обсягами даних без втрати продуктивності. Це важливо для редакцій, які стикаються з різким збільшенням інформаційного потоку, наприклад, під час кризових ситуацій або важливих подій.

Автоматизація також дозволяє уникнути людського фактору, пов'язаного з помилками при ручному створенні рефератів. Система забезпечує стандартизований підхід до реферування, що гарантує стабільну якість матеріалів незалежно від досвіду чи навантаження на редактора [18, с. 108].

Ще одним важливим елементом інтеграції є її здатність до багатомовної обробки текстів. Методика може застосовуватись для реферування новин різними мовами, що розширює можливості використання системи у міжнародних медіа. Інтеграція системи у редакційні процеси також включає налаштування інтерфейсу користувача для зручності роботи редакторів. Гнучкість системи дозволяє легко інтегрувати її у вже існуючі платформи управління контентом. Наприклад, у системах CMS можна створити окремі модулі для автоматичного створення рефератів із збереженням доступу до первинного тексту.

Економічна ефективність впровадження також заслуговує на увагу. Зменшення витрат часу на підготовку рефератів дозволяє скоротити витрати на оплату праці або перерозподілити ресурси на виконання інших важливих завдань.

Таблиця 3.4.

Економічна вигода впровадження автоматизації

Показник	До впровадження	Після впровадження	Зміни
Витрати на оплату праці за місяць	\$15,000	\$10,000	–33%
Додатковий обсяг роботи (тексти)	1,000	5,000	+400%

Впровадження автоматизованої системи також змінює підхід до навчання персоналу. Журналісти та редактори отримують змогу працювати із сучасними технологіями, що підвищує їхню кваліфікацію та залученість до інновацій.

Значний вплив має інтеграція методики на швидкість реагування редакції на актуальні події. У режимі реального часу система здатна обробляти великий обсяг інформації, забезпечуючи швидку підготовку матеріалів для публікації.

Однією з ключових переваг є також підтримка інтерактивного налаштування параметрів реферування. Редактори можуть задавати індивідуальні вимоги до кожного тексту, наприклад, обсяг реферату, акцент на певних аспектах новини чи стиль викладення інформації [19, с. 63].

Таблиця 3.5.

Параметри налаштування реферування

Параметр	Можливі значення	Приклад
Обсяг реферату	10–30% від обсягу тексту	Реферат на 50 слів
Пріоритетні аспекти	Тематика, ключові факти	Політичні події
Стиль викладення	Формальний, інформативний	Інформативний стиль

Інтеграція методики також дозволяє реалізувати принципи прозорості у створенні контенту. Завдяки системам пояснення рішень, користувачі можуть розуміти, чому система обрала певні елементи тексту для реферату.

Автоматизація сприяє також підвищенню залучення аудиторії. Чіткі, лаконічні та інформативні реферати стимулюють читачів переходити до детального перегляду матеріалів, що позитивно впливає на показники відвідуваності ресурсів. Етичний аспект впровадження також заслуговує на увагу. Автоматизована система здатна мінімізувати ризики упередженості, оскільки алгоритми навчання базуються на статистичному аналізі, а не на суб'єктивних оцінках.

Перспективи розвитку методики полягають у її інтеграції з іншими технологіями штучного інтелекту, такими як аналіз тональності тексту чи автоматичне прогнозування популярності новин. Також перспективи розвитку системи включають удосконалення алгоритмів для роботи з мультимедійним контентом, підвищення автономності моделей та інтеграцію з новими джерелами даних. У підсумку, впровадження автоматизованої методики реферування є не лише технічним нововведенням, але й стратегічним кроком до оптимізації роботи редакцій та підвищення ефективності управління контентом. Це рішення дає змогу зосередити увагу редакцій на створенні якісного контенту, залишаючи рутинні завдання на обробку автоматизованими системами [20, с. 147].

Інтеграція методик у редакційні процеси є важливим етапом, що дозволяє досягти високої якості матеріалів, систематизувати та оптимізувати роботу редакторів. Застосування спеціалізованих методик дозволяє не лише покращити кінцевий продукт, але й забезпечити ефективне використання часу та ресурсів.

Програмне забезпечення та технології, що використовуються в редакціях, також повинні бути інтегровані в редакційний процес. Вони дозволяють автоматизувати рутинні завдання, такі як перевірка орфографії, граматики, стилістики, а також забезпечити зручне співробітництво між авторами та редакторами.

Інтеграція методики в редакційний процес передбачає також розробку нових стандартів якості та контролю, що повинні відповідати вимогам сучасного ринку. Це включає не тільки технічні та стилістичні аспекти, але й вимоги до змісту, який повинен бути релевантним, коректним та цікавим для читачів. Редактори повинні забезпечити, щоб всі публікації відповідали високим стандартам якості. Завершальним етапом інтеграції методики є контроль за виконанням всіх норм і стандартів, а також постійне удосконалення редакційних процесів. Це дозволяє створити ефективну систему роботи, яка забезпечує стабільно високу якість матеріалів, що публікуються. Інтеграція інноваційних методик у редакційні процеси є важливим кроком у розвитку медіа та журналістики, що дозволяє не тільки підвищити продуктивність роботи, але й забезпечити більшу точність і якість інформаційного продукту. Впровадження нових технологій і методів роботи стало необхідним у зв'язку з постійними змінами в інформаційному середовищі. У таких умовах методики, що спрощують і автоматизують рутинні завдання, можуть значно змінити характер журналістської роботи, роблячи її більш ефективною та точною.

Одним із ключових аспектів інтеграції методик є здатність до адаптації в залежності від типу медіаорганізації та специфіки контенту. Це дозволяє методиці бути універсальною та застосовуватися не лише в новинних агентствах чи на телебаченні, а й у спеціалізованих видавництвах, що працюють в нішевих сегментах. Враховуючи різноманітність інформаційних потоків, важливо, щоб методика могла підлаштовуватись під кожен конкретний випадок, що робить її вкрай корисною для редакцій будь-якого масштабу.

Процес інтеграції методики автоматизованого реферування в редакційну діяльність також включає численні етапи, починаючи від тестування технології до її впровадження в повсякденну роботу редакції. Цей етап потребує належного планування, оскільки не кожна редакція готова одразу змінити свої робочі процеси. Для успішного інтегрування нової методики необхідно провести ряд навчальних заходів для співробітників, підготувати їх до використання нових інструментів, а також оптимізувати робочі процеси таким чином, щоб технологія

приносила максимальну користь. Реалізація методики автоматизованого реферування дозволяє значно скоротити час, необхідний для підготовки новинного контенту. У випадку з великими обсягами текстів це особливо важливо, оскільки дає можливість швидше реагувати на події, що відбуваються, і зберігати при цьому якість новин. Це також дає змогу зменшити навантаження на журналістів та редакторів, які зазвичай витрачають багато часу на підготовку анотацій або рефератів до новин, що дозволяє зосередитись на більш глибоких і важливих аспектах роботи.

Однією з основних переваг методики є її здатність не тільки автоматизувати процес реферування, але й допомагати виділяти важливі частини інформації, що є найбільш актуальними для аудиторії. Це дозволяє автоматизованій системі працювати не лише як інструмент для створення лаконічних підсумків, але й як інструмент для оцінки значущості різних елементів новини. За допомогою аналізу метаданих можна виявити найбільш важливі факти, підтвердження джерел, а також встановити логічні зв'язки між різними подіями, що може бути надзвичайно корисним для редакторів, які працюють з великими потоками новин. Інтеграція такої методики також включає важливі етичні аспекти. Оскільки в умовах сучасного інформаційного середовища дуже важливо забезпечити достовірність і об'єктивність, методика включає перевірку джерел інформації. Вона дозволяє не тільки автоматично визначати важливі факти, а й здійснювати контроль за правдивістю наданих даних, що в умовах постійного потоку новин є важливим інструментом боротьби з фейковими новинами і маніпуляціями.

Для ефективного функціонування такої методики в редакціях важливо забезпечити тісну взаємодію між автоматизованими системами і людьми, що працюють в команді. Чітке розподілення ролей і функцій, де людина займається більш складними завданнями, а система — рутинною роботою, допомагає забезпечити високу продуктивність і точність. Такі інструменти можуть значно покращити взаємодію між авторами, редакторами та іншими учасниками процесу, а також підвищити рівень співпраці всередині редакції.

Ще одним важливим аспектом є постійне вдосконалення методики на основі зворотного зв'язку. Оскільки інформаційні процеси постійно змінюються, система повинна адаптуватись до нових умов, наприклад, до змін у стилі написання статей або змін у вимогах до форматування інформації. Постійний процес адаптації та навчання є важливою складовою інтеграції цієї методики в редакційну діяльність. Інтеграція автоматизованих методик також дозволяє редакціям більш ефективно реагувати на різноманітні інформаційні запити від громадськості та інших зацікавлених сторін. У разі виникнення кризових ситуацій, коли потрібно швидко розповсюдити важливу інформацію, система автоматизованого реферування може значно прискорити цей процес, забезпечивши високу точність і швидкість при створенні новин. Загалом, впровадження новітніх методик у редакційні процеси є складним, але необхідним етапом у розвитку медіа. Враховуючи сучасні вимоги до якості контенту та швидкості його створення, автоматизація рутинних процесів є важливою складовою вдосконалення діяльності редакцій. Завдяки використанню метаданих та інших сучасних технологій, автоматизовані методики реферування можуть стати невід'ємною частиною роботи редакцій у майбутньому.

Один з головних викликів, які виникають при впровадженні таких методик, – це потреба в постійному розвитку технологій і навичок працівників редакцій. Оскільки технології розвиваються швидкими темпами, редакції повинні бути готові адаптувати нові підходи до роботи, підтримуючи професійну підготовку своїх співробітників.

Розширення можливостей методики, зокрема її адаптація до роботи з мультимедійними форматами, дає великі перспективи для розвитку. У майбутньому методика може бути адаптована для роботи з не тільки текстовими, але й відео- та аудіоматеріалами, що значно розширить її застосування в сучасному медіапросторі. Водночас важливим аспектом є постійний моніторинг і аналіз результатів впровадження нової методики. Тільки через ретельний аналіз результатів можна виявити слабкі місця в системі і вчасно їх виправити. Це забезпечить подальший розвиток технологій і поліпшення якості роботи редакцій.

Таким чином, інтеграція методики автоматизованого реферування в редакційні процеси є важливим кроком у розвитку сучасної журналістики. Цей процес включає не тільки технологічні нововведення, а й культурні зміни в роботі журналістів. Застосування такої методики дозволяє не тільки полегшити роботу редакцій, а й забезпечити більшу точність, швидкість та прозорість в інформаційному просторі. У перспективі можна очікувати подальше вдосконалення цих методик та їх більш широке застосування в різних медіасекторів.

Інтеграція методик у редакційні процеси є комплексним і багатоетапним завданням, яке потребує постійного вдосконалення та адаптації до змін у технологіях, вимогах ринку та потребах читачів. Ключовим елементом є не лише технічні інструменти, а й розуміння важливості професійного розвитку та комунікації між усіма учасниками процесу. Інтеграція методик реферування у редакційні процеси демонструє значне підвищення продуктивності та зниження витрат часу. Незважаючи на деякі виклики, зокрема, потребу у налаштуванні системи під специфічні стилістичні вимоги, результати свідчать про високу ефективність впровадження. Це відкриває перспективи для масштабного використання методики в редакціях новинних агентств, а також її подальшого вдосконалення.

3.2. Розвиток та адаптація методики

Розвиток методики реферування новинних текстів передбачає її поступове вдосконалення та адаптацію до нових вимог сучасного інформаційного простору. Цей процес включає технологічне оновлення, інтеграцію з інноваційними рішеннями та врахування змін у суспільному попиті на інформацію.

Одним із ключових напрямків розвитку є підвищення точності алгоритмів семантичного аналізу. Для цього використовуються сучасні технології глибокого навчання, зокрема архітектури трансформерів, які забезпечують більш точне розуміння смислових зв'язків у текстах. Адаптація методики до специфічних

тематичних галузей є ще одним пріоритетним завданням. Новинні тексти можуть охоплювати широкий спектр тем, від політики до розваг. Кожна з них має свої стилістичні особливості, які необхідно враховувати при розробці рефератів. Методика також має враховувати динаміку зміни формату новин. Наприклад, зростає популярність коротких інформаційних матеріалів, які вимагають створення ще більш стислих та лаконічних рефератів. Це створює нові виклики для адаптації системи [21, с. 141].

Інтеграція мультимедійних форматів у процес реферування є перспективним напрямком. Наприклад, відео чи аудіоматеріали можуть автоматично перетворюватися в текстові реферати, що значно розширює функціональність методики. Ще одним важливим аспектом розвитку є впровадження адаптивного навчання. Система має враховувати зворотний зв'язок від редакторів і користувачів, що дозволяє їй поступово вдосконалювати свої алгоритми на основі реальних даних.

Розширення можливостей методики також включає її інтеграцію з аналітичними системами. Наприклад, алгоритми можуть аналізувати вплив кожного реферату на поведінку аудиторії, допомагаючи редакціям оцінювати ефективність новинних матеріалів. Зростання кількості новинних платформ і каналів комунікації створює потребу у забезпеченні масштабованості методики. Вона повинна працювати ефективно не лише для традиційних новинних агентств, але й для соціальних мереж та незалежних блогів [22, с. 31].

Адаптація до нових мов та культурних контекстів також є важливим аспектом розвитку. Це дозволить використовувати методику у глобальному масштабі, забезпечуючи якісний аналіз і реферування для аудиторій з різних країн. Етичний розвиток методики включає врахування ризиків маніпуляцій та дезінформації. Система має мати вбудовані механізми перевірки достовірності джерел та уникнення упередженості у створенні рефератів.

Технологічне оновлення передбачає використання сучасних обчислювальних платформ, які забезпечують швидку обробку великих обсягів

даних. Хмарні технології дозволяють масштабувати систему залежно від поточного навантаження.

Важливим напрямком адаптації є врахування змін у медіаспоживанні аудиторії. Сьогодні споживачі віддають перевагу персоналізованому контенту, тому методика має включати механізми налаштування рефератів під індивідуальні потреби користувачів. Методика повинна забезпечувати автоматизацію рутинних завдань редакцій, залишаючи більше часу для творчої роботи. Це досягається завдяки впровадженню інтелектуальних інструментів, які спрощують роботу з текстами. Адаптація також передбачає можливість роботи з новими форматами даних, наприклад, інтерактивними чи гейміфікованими новинами. Система має бути достатньо гнучкою для інтеграції таких форматів у процес реферування.

Зростання кількості даних створює виклики для зберігання та обробки інформації. Методика повинна бути оптимізована для роботи з великими даними, забезпечуючи високу продуктивність навіть у складних умовах. Розвиток системи включає розширення функціоналу для роботи з різними типами текстів. Наприклад, методика може використовуватись не лише для новин, але й для академічних статей, маркетингових матеріалів чи інших видів контенту [23, с. 87].

Ще одним важливим аспектом є забезпечення конфіденційності даних. Система має дотримуватися всіх сучасних стандартів захисту інформації, щоб уникнути витоку даних або їх несанкціонованого використання. Залучення штучного інтелекту дозволяє створювати реферати, які виглядають максимально природно, майже не відрізняючись від створених людиною. Це підвищує довіру користувачів до автоматизованих рішень.

Розвиток методики включає також її застосування у кризових ситуаціях. Наприклад, під час надзвичайних подій система може забезпечувати швидке створення рефератів для оперативного інформування аудиторії.

Інтеграція з іншими технологіями, такими як аналіз емоційного забарвлення тексту, дозволить створювати реферати, які враховують не лише факти, але й емоційну складову повідомлення. Методика має також враховувати специфіку

різних медійних платформ. Наприклад, реферати для соціальних мереж мають бути більш стислими, ніж для традиційних новинних сайтів. Система має розвиватися у напрямку підтримки голосового вводу та виводу. Це дозволить створювати голосові реферати, що зручно для користувачів, які споживають інформацію через голосові асистенти.

Методика може інтегруватись із системами автоматизованого перекладу, що дозволить створювати багатомовні реферати без втрати якості. Це сприятиме популяризації контенту у глобальному масштабі. Оцінка ефективності впровадження повинна проводитися регулярно, використовуючи метрики, які враховують не лише кількісні, але й якісні показники. Сучасний розвиток технологій штучного інтелекту відкриває нові горизонти для вдосконалення методики. Завдяки інноваціям вона може стати універсальним інструментом для аналізу, синтезу та реферування текстів у різних галузях [24, с. 366].

Подальший розвиток методики також повинен включати інтеграцію елементів когнітивної лінгвістики для глибшого розуміння не лише структурних, але й концептуальних аспектів тексту. Це дозволить автоматизованій системі враховувати складні взаємозв'язки між термінами, концептами та їхнім культурним контекстом. Особливу увагу слід приділити створенню методик адаптації до нових стилістичних форматів тексту, які постійно з'являються в журналістиці. Наприклад, сучасні медіа дедалі частіше використовують змішані жанри, де поєднуються елементи новин, аналітики й оповідання.

Актуальним напрямком є вдосконалення алгоритмів виявлення тональності тексту. Це дозволить системі не лише створювати об'єктивні реферати, але й формувати стислий огляд тексту, враховуючи його емоційне забарвлення, що важливо для різних аудиторій. Інтеграція віртуальних асистентів також відкриває нові можливості для адаптації методики. Наприклад, система може автоматично генерувати голосові реферати новин, що буде зручним для користувачів під час керування автомобілем чи виконання інших завдань. Розширення використання хмарних обчислень дозволить зробити методику доступнішою для невеликих

редакцій. Хмарні сервіси знизять витрати на обладнання та забезпечать високу продуктивність навіть при обробці великих обсягів даних [25, с. 310].

Додатковим аспектом розвитку є інтеграція з платформами соціальних мереж. Автоматизована система може створювати адаптовані реферати для публікацій у різних соціальних каналах, враховуючи особливості їхньої аудиторії та форматів контенту. Покращення інтерфейсу користувача в редакційних системах, що використовують методику, дозволить зробити її більш зручною та інтуїтивно зрозумілою. Це сприятиме зменшенню часу, який витрачається на навчання співробітників.

Використання даних зворотного зв'язку від аудиторії також відіграватиме важливу роль у вдосконаленні методики. Система може аналізувати, як користувачі взаємодіють із рефератами, і на основі цього покращувати алгоритми [26, с. 265].

Іншим напрямком є вдосконалення механізмів боротьби з фейковою інформацією. Наприклад, система може аналізувати джерела даних і включати до рефератів лише ті, які відповідають критеріям достовірності. Нарешті, перспективним є розвиток інтегрованих систем, які поєднують автоматичне реферування з функціями прогнозування. Це дозволить створювати не лише реферати, а й аналітичні висновки, що можуть стати основою для стратегічних рішень у медіа чи інших галузях [27, с. 66].

Розвиток і адаптація методики є динамічним процесом, що дозволяє їй відповідати сучасним вимогам інформаційного простору. Це сприяє створенню якісних, достовірних і ефективних рефератів, які задовольняють потреби як редакцій, так і кінцевих користувачів.

Одним із ключових напрямків розвитку є впровадження адаптивних алгоритмів, які здатні враховувати специфічні особливості текстів залежно від контексту. Це дозволяє значно покращити ефективність роботи систем реферування, забезпечуючи точніше розпізнавання ключових елементів тексту та адаптацію методики до змін у структурі новинних статей. Наприклад, методики, що використовують трансформерні архітектури (такі як BERT чи GPT), уже

демонструють високий рівень адаптивності та можуть працювати з текстами різних тематик і жанрів.

Розвиток методики також спрямований на інтеграцію нових джерел даних, таких як мультимедійні матеріали. Це дозволяє враховувати не лише текстову, але й візуальну чи аудіоінформацію, яка доповнює основний зміст новинного повідомлення. Наприклад, графічні дані, діаграми або відеофрагменти, які супроводжують текст, можуть бути використані для створення більш повного уявлення про контекст новини. Ще одним перспективним напрямком є використання онтологічних моделей для покращення семантичного аналізу текстів. Онтології дозволяють структурувати знання про предметну область, що суттєво підвищує здатність системи розуміти взаємозв'язки між окремими елементами тексту. Це, у свою чергу, дозволяє створювати реферати, які більш точно відображають зміст початкових документів.

Важливою складовою адаптації методики є врахування багатомовності текстів. Сучасні системи реферування повинні забезпечувати однаково високу якість роботи незалежно від мови, якою написано текст. Для цього використовуються багатомовні моделі, які здатні навчатися на текстах різними мовами, а також алгоритми перекладу, що дозволяють інтегрувати тексти різних мов у єдину систему аналізу. Адаптація до вимог користувачів також є пріоритетним завданням у розвитку методик реферування. Користувачі часто мають специфічні запити щодо формату, стилю або рівня деталізації рефератів. Тому сучасні системи реферування все частіше включають механізми персоналізації, які дозволяють налаштовувати результат відповідно до індивідуальних потреб користувача.

Іншою важливою тенденцією є інтеграція методик реферування з іншими інформаційними системами, такими як системи управління знаннями, CRM або аналітичні платформи. Це дозволяє розширити функціональність методики, забезпечуючи її використання в більш широкому спектрі завдань. Наприклад, результати реферування можуть бути автоматично передані в систему для подальшого аналізу або прийняття рішень. Суттєвим фактором у розвитку

методики є забезпечення її масштабованості. Зі збільшенням обсягів даних, що аналізуються, системи реферування повинні залишатися ефективними та швидкими. Для цього використовуються хмарні технології та розподілені обчислення, які дозволяють обробляти великі масиви текстів у режимі реального часу.

Безпека даних також є важливим аспектом у розвитку методик реферування. Забезпечення конфіденційності інформації, що обробляється, стає ключовою вимогою, особливо у випадках роботи з чутливими даними. Для цього впроваджуються технології шифрування, а також розробляються алгоритми, які мінімізують ризики витоку інформації.

Окремо слід зазначити роль етичних аспектів у розвитку методик реферування. Системи повинні враховувати можливість упереджень у текстах та запобігати їхньому посиленню в процесі реферування. Для цього використовуються алгоритми перевірки достовірності інформації та механізми виявлення маніпулятивних елементів у текстах. Значний внесок у розвиток методик реферування роблять технології візуалізації. Графічне представлення результатів реферування дозволяє зробити інформацію більш зрозумілою та зручною для користувача. Наприклад, використання інфографіки або інтерактивних діаграм може значно підвищити якість сприйняття рефератів.

У рамках адаптації методики активно досліджуються можливості використання даних про поведінку користувачів для вдосконалення алгоритмів реферування. Аналіз таких даних дозволяє краще зрозуміти, які саме аспекти тексту є найбільш важливими для користувачів, і відповідним чином адаптувати процес реферування.

Розвиток та адаптація методики також охоплюють створення інструментів для зворотного зв'язку від користувачів. Такі інструменти дозволяють збирати відгуки про якість рефератів і використовувати їх для подальшого вдосконалення системи. Це сприяє створенню методик, які максимально відповідають потребам користувачів.

Одним із ключових викликів у розвитку методик є забезпечення їхньої інтеграції в освітні процеси. Автоматизовані системи реферування можуть бути корисними для студентів і викладачів, забезпечуючи швидкий доступ до короткого викладу змісту наукових праць або навчальних матеріалів. Для цього методики адаптуються до специфіки академічних текстів і вимог освітніх програм. Подальший розвиток методик реферування залежить від міждисциплінарного підходу до їхнього вдосконалення. Інтеграція знань із різних галузей, таких як лінгвістика, інформатика, психологія та соціологія, дозволяє створювати більш досконалі моделі, які враховують усі аспекти процесу реферування.

У перспективі очікується поява нових інструментів для оцінки якості реферування. Такі інструменти будуть базуватися на об'єктивних метриках, які дозволяють вимірювати точність, інформативність та релевантність рефератів. Це сприятиме підвищенню стандартів якості у цій сфері. Розвиток та адаптація методики реферування текстів є важливим завданням, що потребує комплексного підходу. Завдяки впровадженню нових технологій, адаптації до вимог користувачів і забезпеченню високої якості роботи, сучасні методики реферування стають незамінним інструментом для аналізу великих обсягів текстової інформації та забезпечення ефективної комунікації в умовах інформаційного суспільства. Розвиток та адаптація методики є важливим етапом для забезпечення її ефективного функціонування в умовах постійних змін, що відбуваються в медіапросторі. У світі, де інформація постійно змінюється, а технології еволюціонують з неймовірною швидкістю, необхідно постійно вдосконалювати методику, щоб вона могла відповідати новим вимогам і викликам. Однак, щоб досягти цього, потрібно враховувати кілька ключових аспектів: інтеграцію нових технологій, адаптацію до змін у журналістській практиці, а також постійну оптимізацію самих алгоритмів методики.

Одним із головних елементів адаптації є поєднання традиційних методів журналістики з новими технологічними інструментами. В умовах зростаючої конкуренції на інформаційному ринку медіа повинні вчасно реагувати на зміни,

які відбуваються в споживчих перевагах та технологічних трендах. З цієї точки зору методика автоматизованого реферування повинна постійно вдосконалюватися для забезпечення зручності користувачів і підвищення якості результатів.

Одним із найбільших викликів на цьому етапі є необхідність інтеграції нових інструментів з уже існуючими технологіями, які використовуються в редакційних процесах. У багатьох редакціях досі застосовуються традиційні методи обробки інформації, такі як ручне створення анотацій або рефератів до новин, що займає багато часу і зусиль. Однак із розвитком штучного інтелекту та машинного навчання стає можливим використання більш складних і точних алгоритмів для автоматизації цих процесів. Розвиток і адаптація методики повинні включати в себе вдосконалення таких алгоритмів, щоб вони могли більш точно виділяти важливі моменти з новинних статей, забезпечуючи високий рівень точності та коректності.

Технологічний розвиток не обмежується лише вдосконаленням алгоритмів. У світлі постійного розвитку мультимедійних форматів, таких як відео, аудіо та інфографіка, методика повинна бути адаптована до роботи з такими форматами. Наприклад, автоматизоване реферування повинно бути здатне працювати не лише з текстовими матеріалами, а й з відео та аудіо записами, аналізуючи зміст відеоматеріалів та створюючи точні реферати на основі цього контенту. Такий підхід значно розширює можливості методики та дозволяє використовувати її не тільки для текстових новин, а й для більш складних мультимедійних проектів.

Адаптація методики до мультимедійних форматів вимагає вдосконалення алгоритмів, що забезпечують автоматичне розпізнавання ключових моментів у відео- та аудіоформатах. Для цього необхідно використовувати технології комп'ютерного зору та обробки природної мови, які дозволяють аналізувати зображення, звук та текст одночасно. Застосування таких підходів допоможе створювати ще більш точні та інформативні реферати, що враховують різноманітні аспекти контенту.

Ще одним важливим напрямком розвитку є вдосконалення алгоритмів перевірки джерел і достовірності інформації. В умовах інформаційної епідемії фейкових новин, що стала характерною для сучасного медіапростору, дуже важливо, щоб автоматизовані системи не тільки створювали реферати, а й могли перевіряти інформацію на достовірність. Методика повинна включати механізми, що автоматично визначають джерела інформації, порівнюють їх з перевіреними базами даних і оцінюють рівень достовірності новини. Для цього необхідно інтегрувати системи перевірки фактів, які за допомогою штучного інтелекту можуть оперативно виявляти маніпуляції або недостовірну інформацію.

Одним із способів розвитку методики є також використання зворотного зв'язку від користувачів системи. Це важливо для подальшої адаптації методики до конкретних вимог редакцій, оскільки різні медіаорганізації можуть мати свої специфічні вимоги щодо того, як мають виглядати готові реферати або які аспекти новин потрібно акцентувати. Використання зворотного зв'язку допомагає зібрати дані щодо ефективності методики та її застосування в реальних умовах, що в свою чергу дозволяє постійно вдосконалювати систему і робити її більш ефективною.

Адаптація методики також включає вдосконалення її функціоналу, що забезпечує більшу гнучкість і здатність до інтеграції з іншими редакційними інструментами. Наприклад, вона повинна бути сумісною з сучасними системами управління контентом (CMS) і іншими інструментами для автоматизації робочих процесів, що використовуються редакціями. Таке поєднання допомагає створювати єдину інфраструктуру для обробки новинного контенту, де кожен етап — від збору інформації до її публікації — автоматизований і організований на високому рівні.

Крім того, важливим аспектом є оптимізація алгоритмів методики для різних мовних та культурних контекстів. У глобалізованому світі медіа, що функціонують в багатьох країнах і працюють з різними аудиторіями, необхідно враховувати мовні та культурні відмінності. Алгоритми повинні бути здатні правильно обробляти тексти, написані на різних мовах, і коректно інтерпретувати культурні

контексти, що відображаються в новинах. Це вимагає розробки адаптивних моделей, які можуть враховувати лексичні, синтаксичні та семантичні особливості кожної мови.

Одним із важливих аспектів розвитку методики є також інтеграція її з соціальними мережами та іншими сучасними цифровими платформами. Багато новин та інформації сьогодні поширюються через соціальні мережі, де кожен користувач може бути як споживачем, так і творцем контенту. Для того, щоб методика була ефективною в цих умовах, вона повинна вміти обробляти контент з різних джерел, включаючи соціальні мережі, блоги та інші онлайн-платформи. Це дозволить журналістам автоматично отримувати реферати новин не лише з традиційних медіа, але й з різних інтернет-ресурсів.

Підсумовуючи, розвиток та адаптація методики автоматизованого реферування є постійним процесом, що включає вдосконалення як технологій, так і самого підходу до роботи з інформацією. Адаптація до нових умов медіапростору, інтеграція з новими платформами та технологіями, а також постійне вдосконалення алгоритмів є ключовими елементами розвитку цієї методики. У майбутньому, з урахуванням швидкого розвитку технологій, таких як штучний інтелект та машинне навчання, можна очікувати подальше покращення автоматизованих методик, що дозволить ще більше підвищити ефективність роботи редакцій та зробити їх більш гнучкими й здатними до адаптації до нових умов. Розвиток і адаптація методики є динамічним процесом, що дозволяє їй відповідати сучасним вимогам інформаційного простору. Це сприятиме створенню якісних, достовірних і ефективних рефератів, які задовольняють потреби як редакцій, так і кінцевих користувачів.

ВИСНОВКИ

У результаті дослідження була розроблена методика автоматизованого реферування новинних статей, яка враховує сучасні виклики інформаційного середовища. Її основою стали метадані, що забезпечують структурований підхід до аналізу та синтезу текстів.

Однією з ключових переваг методики є її адаптивність. Вона може застосовуватись до різних тематичних напрямків і типів текстів, що робить її універсальним інструментом для редакцій різного масштабу. Під час реалізації прототипу було підтверджено ефективність використання метаданих для виділення ключових ідей та створення лаконічних рефератів. Це дозволяє значно скоротити час на обробку інформації та зберегти основний зміст повідомлення.

Інтеграція методики у редакційні процеси сприяє підвищенню продуктивності роботи журналістів. Вона дозволяє автоматизувати рутинні завдання, звільняючи час для творчої та аналітичної діяльності.

Система також враховує етичні аспекти, забезпечуючи перевірку достовірності джерел та мінімізацію ризику маніпуляцій. Це особливо важливо в умовах зростаючої кількості фейкової інформації в медіа. Розроблена методика має значний потенціал для подальшого розвитку. Її адаптація до мультимедійних форматів, інтеграція з соціальними платформами та вдосконалення алгоритмів дозволять зробити її ще більш ефективною та затребуваною. У результаті розробки інноваційної методики автоматизованого реферування новинних статей було досягнуто значного прогресу в оптимізації процесів обробки інформації та підвищенні ефективності роботи редакцій. Враховуючи вимоги сучасного інформаційного середовища, ця методика виявилася важливим інструментом для швидкої та якісної обробки великих обсягів текстів, що є необхідним в умовах динамічного розвитку медіа та численних інформаційних потоків.

Однією з основних переваг запропонованої методики є її здатність адаптуватися до різних контекстів та форм форматів. Можливість застосування

алгоритмів, що ґрунтуються на метаданих, дозволяє автоматизувати виділення найбільш важливих аспектів кожної новини без втрати контексту. Це дає змогу створювати чіткі та зрозумілі реферати, що мають високу інформативну цінність і відповідають стандартам журналістики. Технологія, яка активно використовує структуровані дані, дає можливість точно і швидко виділяти основні моменти, тим самим значно скорочуючи час, необхідний на ручне редагування та створення анотацій.

Завдяки метаданим методика забезпечує не лише покращення швидкості обробки, а й точність виділення найбільш важливих елементів новини, таких як ключові факти, дані, цитати та підтвердження джерел. Вона дозволяє журналістам зосереджуватися на аналізі нових тенденцій, пошуку важливих зв'язків між подіями та інтерпретації інформації. Таким чином, автоматизація рутинних завдань, таких як виділення головних ідей статей, надає можливість для розвитку більш творчих та стратегічних аспектів журналістської роботи.

Інтеграція цієї методики в редакційні процеси стає важливим етапом у розвитку сучасної журналістики, адже вона дозволяє не лише підвищити продуктивність праці журналістів, а й сприяти кращому організуванню інформаційних потоків. В умовах великої кількості новин, що постійно з'являються, важливим є вміння швидко і точно оцінити їх значущість. Автоматизоване реферування дозволяє редакціям скорочувати час на обробку новин, що є особливо важливим у процесі оперативного реагування на актуальні події. Це дає змогу не тільки швидше публікувати матеріали, але й забезпечити їх більшу точність та узгодженість.

Враховуючи особливості сучасного медіапростору, методика автоматизованого реферування також має великий потенціал для боротьби з фейковою інформацією. Інтеграція алгоритмів, які здійснюють перевірку достовірності джерел, дозволяє мінімізувати ризик маніпуляцій та сприяти створенню більш прозорого інформаційного середовища. Потрібно зазначити, що етичні аспекти є важливим компонентом цієї методики, адже вони забезпечують не лише правильність переданих фактів, але й довіру аудиторії до публікацій.

Інноваційна методика враховує зростаючий виклик боротьби з фейковими новинами та маніпуляціями. У світі, де штучно створена інформація може поширюватися дуже швидко, важливо мати інструменти, які можуть допомогти вчасно ідентифікувати та нейтралізувати таку інформацію. Зокрема, методика дозволяє здійснювати перевірку фактів та джерел, а також виявляти спроби маніпулювання через неправдиві або спотворені твердження, що можуть бути частиною новин. Автоматизоване реферування вимагає інтеграції алгоритмів, що здатні здійснювати комплексну перевірку джерел та аналізувати контекст новини для виявлення можливих маніпуляцій або неточностей. Розробка цієї методики також відкриває нові можливості для адаптації до мультимедійних форматів. У наш час новини часто подаються в супроводі відео, аудіо та графічних елементів, що вимагає від системи здатності працювати не тільки з текстом, а й з іншими типами даних. Перспективи інтеграції методики в мультимедійне середовище включають автоматизоване реферування не лише текстових статей, але й відео- та аудіоконтенту. Застосування метаданих до цих форматів дозволяє здійснювати більш комплексний аналіз і створювати лаконічні реферати, що відображають основний зміст поданого матеріалу.

Враховуючи можливості розвитку штучного інтелекту та машинного навчання, алгоритми для автоматизованого реферування мають великий потенціал для вдосконалення. Машинне навчання дозволяє системі вчитися на нових даних, адаптуючи свої алгоритми до змін у стилі подачі новин та нових вимог журналістики. Згодом це може дозволити досягти ще більшої точності у визначенні головних ідей та фактів, що забезпечить більш високу якість кінцевого продукту. Окремо варто зазначити, що інтеграція цієї методики в редакційні процеси не лише покращує оперативність роботи редакцій, але й сприяє підвищенню рівня журналістської етики. Автоматизація процесу реферування дає можливість не тільки скоротити час, необхідний для обробки новин, але й зменшити ймовірність помилок, що можуть виникнути під час ручної обробки. Це важливо в умовах, коли точність і достовірність інформації є критичними для створення якісного та надійного контенту.

Впровадження цієї методики також створює можливості для більш детального аналізу тенденцій у медіапросторі, оскільки автоматизоване реферування дозволяє швидко обробляти великі обсяги даних і знаходити важливі взаємозв'язки між різними новинами та подіями. Це особливо актуально для дослідників медіа, які прагнуть зрозуміти загальні тенденції у інформаційному просторі. Таким чином, автоматизоване реферування новинних статей, засноване на використанні метаданих, є важливим кроком до створення більш ефективних і продуктивних редакційних процесів. Враховуючи численні переваги цієї методики, можна очікувати, що вона стане основним інструментом у роботі багатьох редакцій і медіаорганізацій, сприяючи підвищенню якості журналістики, зниженню ризиків маніпулювання інформацією та розвитку інноваційних підходів до обробки новин. Перспективи подальшого розвитку методики включають розширення її застосування до нових форматів інформації та інтеграцію з іншими сучасними технологіями, такими як штучний інтелект, великий аналіз даних і нові медіаплатформи.

У підсумку, автоматизоване реферування новинних статей є важливим напрямком розвитку сучасної журналістики. Використання запропонованої методики не лише підвищує якість інформаційного продукту, але й сприяє формуванню прозорого та достовірного інформаційного простору.

Результати дослідження апробовано шляхом публікації тез доповідей на конференціях:

1. Гончарова А.А., Золотухіна О.А., Огляд інформаційно-аналітичної діяльності в контексті покращення процесу пошуку інформації на веб-ресурсах та методів реферування статей за допомогою метаданих // Всеукраїнська науково-технічна конференція «Застосування програмного забезпечення в ІКТ». – Київ: ДУІКТ, 2024. – С.166-169.
2. Гончарова А.А., Золотухіна О.А., Аналіз методів реферування новинних статей на основі метаданих за допомогою штучного інтелекту // XXIV Міжнародна науково-практична конференція «Інформаційні технології та безпека» (ІТС-2024). – Київ: НАН України, 2024. – *подано до друку*.

ПЕРЕЛІК ПОСИЛАНЬ

1. Мазурець О.В., Кліменко В.І., Живілік А.В. Аналіз сучасних методів автоматизації анотування та реферування текстів. 2015. 4 с. URL: <https://elar.khmnu.edu.ua/bitstream/123456789/6816/1/060.pdf>
2. Філіпова, Л. Я.; Захарова, І. В. Аналітична складова інформаційної діяльності: уточнення сутності, ознак і процесів. Вісник Харківської державної академії культури, 2009, 28: 44-52. URL: http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?I21
3. Волохід О.М. Каталогізація цифрових ресурсів Інтернет. Дублінське ядро метаданих: Посібник. Кіровоград, 2003. 72с. URL: https://old.library.kr.ua/dc/DC_Volokhin.pdf
4. Дибіна А.В., Лазаренко О.В. Побудова текстової бази в системі автоматичного реферування на основі структурно-семантичного аналізу тексту. Проблеми семантики слова, речення та тексту. №26. 2011. С. 105-111. URL: <http://surl.li/poqclr>
5. Крайовський В.А., Литвин В.В., Шаховська Н.Б. Основні підходи до розроблення програмного комплексу автоматичного реферування текстових документів. Збірник наукових праць Інституту проблем моделювання в енергетиці ім. ГЄ Пухова НАН України. 2009. 9 с. URL: <http://surl.li/myrnbi>
6. Математичне та імітаційне моделювання систем. МОДС 2018: тези доповідей Тринадцятої міжнародної науково-практичної конференції. ЧНТУ. Чернігів. 2018. 392 с. URL: <http://surl.li/iffgtd>
7. Огнівчук Л.М., Михайлюк А.Ю. Побудова системи автоматичного реферування тексту на основі ієрархічної нечіткої нейронної мережі. Збірник матеріалів І Української конференції молодих науковців "Інформаційні технології-2014". 2014. С. 164-165. URL: <http://surl.li/cwudqj>
8. Михайлюк А.Ю., Михайлюк О.С., Огнівчук Л.М., Тарасенко В.П. Організація системи автоматичного реферування тексту на основі нечіткої ієрархічної нейронної мережі. Радіоелектронні і комп'ютерні системи. №4. 2016. С. 53-68. URL: <http://surl.li/srthdc>

9. Каніщева О.В. Використання карт відношень (TRM) для автоматичного реферування. Вісник Національного університету Львівська політехніка. Інформаційні системи та мережі. №770. 2013. С. 108-122. URL: <http://surl.li/piqusa>
10. Романюк А. Сучасний стан і тенденції розвитку систем автоматичного реферування. Наукові праці Національної бібліотеки України ім. ВІ Вернадського. №37. 2013. С. 253-264. URL: <http://surl.li/oiuqmc>
11. Гончарук Є.О., Діброва Я.І., Коломієць Н.О. Огляд сучасних досліджень у галузі автоматичного реферування. Наукові дослідження–інструмент вирішення актуальних проблем економіки. №6. 6 с. URL: <http://surl.li/ecscvs>
12. Бородіна О.О., Зозуля М.О. Особливості автоматизованого реферування текстів. Чернігівський національний технологічний університет. 2018. URL: <http://surl.li/kmqvey>
13. Шкляревський В.Г. Сучасні корпуси текстів: вимоги до метаданих. Науковий вісник кафедри ЮНЕСКО Київського національного лінгвістичного університету. Філологія, педагогіка, психологія. №30. 2015. С. 241-246. URL: <http://surl.li/yvtuck>
14. Захарова О.В. Класифікація метаданих великих даних. Проблеми програмування. 2019. С. 53-74. URL: <http://surl.li/prjfyu>
15. Косяченко С.В. Каталогізація та ретрокаталогізація газетних видань (на прикладі Національної бібліотеки України імені Ярослава Мудрого). 2020. URL: <http://surl.li/misfbl>
16. Гладун А.Я., Рогушина Ю.В. Семантичні технології: принципи та практики. ТОВ «ВД «АДЕФ-Україна». Київ. 2016. 308 с. URL: <http://surl.li/cbrlzs>
17. Михайлюк А.Ю., Михайлюк О.С., Огнівчук Л.М., Тарасенко В.П. Організація системи автоматичного реферування тексту на основі нечіткої ієрархічної нейронної мережі. Радіоелектронні і комп'ютерні системи. №4. 2016. С. 53-68. URL: <http://surl.li/edasic>
18. Каніщева О.В. Використання карт відношень (TRM) для автоматичного реферування. Вісник Національного університету Львівська

політехніка. Інформаційні системи та мережі. №770. 2013. С. 108-122. URL: <http://surl.li/lfxlrt>

19. Шаховська Н., Литвин В., Крайовський В. Інформаційна система реферування множини документів, поданих у різних форматах, базована на онтології. Вісник Національного університету «Львівська Політехніка». №672. 2010. С. 63-71. URL: <http://surl.li/clysxp>

20. Бармак О.В., Мазурець О.В., Живілік А.В. Інформаційна технологія автоматизованого анотування та реферування цифрових текстів. Вісник Хмельницького національного університету. Технічні науки. №4. 2017. С. 147-158. URL: <http://surl.li/nhadyu>

21. Заболотня Т.М., Федченко Н.В. Модифікований метод автоматичного реферування текстів з використанням тематично зв'язаного ранжування речень. Проблеми інформаційних технологій. №1. 2016. С. 141-147. URL: <http://surl.li/jjbgca>

22. Коханова І. Проблеми та похибки методів автоматизованого реферування документів. Вісник Книжкової палати. №9. 2014. С. 31-32. URL: <http://surl.li/ranlew>

23. Бісікало О.В., Богач І.В. Складність класу семантико-залежних задач обробки тексту. Вісник Вінницького політехнічного інституту. №3. 2016. С. 87-93. URL: <http://surl.li/paxshj>

24. Олійник Ю.О. Деякі аспекти аналізу потоків текстових даних. Міжнародна науково-технічна конференція Інформаційні технології в металургії та машинобудуванні. 2020. С. 366-369. URL: <http://surl.li/fiwnsh>

25. Robertson C.E., Harris J.K., Wagner B.D., Granger D., Browne K., Tatem B., Frank D.N. Explicit: graphical user interface software for metadata-driven management, analysis and visualization of microbiome data. Bioinformatics. №29(23). 2013. С. 3100-3101. URL: <http://surl.li/xxgaxc>

26. Bui D.D.A., Del Fiore G., Hurdle J.F., Jonnalagadda S. Extractive text summarization system to aid data extraction from full text in systematic review development. Journal of biomedical informatics. №64. 2016. С. 265-272. URL: <http://surl.li/xyzexq>

27. Gambhir M., Gupta V. Recent automatic text summarization techniques: a survey. *Artificial Intelligence Review*. №47(1). 2017. C. 1-66. URL: <http://surl.li/ggpagf>
28. Liu S., Zhou M.X., Pan S., Song Y., Qian W., Cai W., Lian X. Tiara: Interactive, topic-based visual text summarization and analysis. *ACM Transactions on Intelligent Systems and Technology*. №3(2). 2012. C. 1-28. URL: <http://surl.li/ralyrn>
29. Ahuja N., Bansal R., Ingram W.A., Jude P., Kahu S., Wang X. Big Data Text Summarization: Using Deep Learning to Summarize Theses and Dissertations. 2018. URL: <http://surl.li/masbqd>
30. Li L., Geissinger J., Ingram W.A., Fox E.A. Teaching natural language processing through big data text summarization with problem-based learning. *Data and Information Management*. №4(1). 2020. C. 18-43. URL: <http://surl.li/ldwzgm>
31. Safder I., Hassan S.U., Visvizi A., Noraset T., Nawaz R., Tuarob S. Deep learning-based extraction of algorithmic metadata in full-text scholarly documents. *Information processing & management*. №57(6). 2020. URL: <http://surl.li/xpvnfk>
32. Mlynarski A. Automatic text summarization in digital libraries. University of Lethbridge. 2006. URL: <http://surl.li/paague>
33. Koniaris M., Galanis D., Giannini E., Tsanakas P. Evaluation of automatic legal text summarization techniques for greek case law. *Information*. №14(4). 2023. C. 250. URL: <http://surl.li/txzzfb>
34. Tejasree S., Naseera S. Improved Clustering Technique Using Metadata for Text Mining. *Innovations in Computer Science and Engineering: Proceedings of the Sixth ICICSE 2018*. Springer Singapore. 2019. C. 243-250. URL: <http://surl.li/pfyaff>
35. Sapra S.J., Thakur S.A., Kapse A.S. Constructive approach for text summarization using advanced techniques of deep learning. *Intelligent Data Communication Technologies and Internet of Things: Proceedings of ICICI 2021*. Springer Nature Singapore. 2022. C. 895-904. URL: https://link.springer.com/chapter/10.1007/978-981-16-7610-9_65

ДОДАТОК А. ДЕМОНСТРАЦІЙНІ МАТЕРІАЛИ



ДЕРЖАВНИЙ УНІВЕРСИТЕТ ІНФОРМАЦІЙНО-
КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ

НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ІНФОРМАЦІЙНИХ
ТЕХНОЛОГІЙ

КАФЕДРА ІНЖЕНЕРІЇ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ



Магістерська робота

«МЕТОДИКА РЕФЕРУВАННЯ НОВИННИХ СТАТЕЙ НА ОСНОВІ МЕТАДАНИХ»

Виконала: студентка групи ПДМ-62 Гончарова Анна Андріївна

Керівник: к.т.н., доц., доцент кафедри ІПЗ Золотухіна Оксана
Анатоліївна

Київ - 2025

МЕТА, ОБ'ЄКТ ТА ПРЕДМЕТ ДОСЛІДЖЕННЯ

Мета роботи: підвищення якості обробки та систематизації новинного контенту за рахунок автоматизованого реферування новинних статей на основі метаданих.

Об'єкт дослідження: процес автоматизованого реферування новинних статей.

Предмет дослідження: методи, моделі та алгоритми автоматизованого реферування новинних статей.

АКТУАЛЬНІСТЬ РОБОТИ

Зростання контенту

Кількість новинного контенту невідмінно зростає, що ускладнює його обробку та аналіз.

Потреба в автоматизації

Автоматизація обробки новин є ключовим фактором для ефективного аналізу величезного потоку інформації.

Зростаюча потреба в автоматизації обробки новинного контенту зумовлена не лише його обсягом, але й швидкістю появи нової інформації. Важливо зберігати баланс між швидкістю обробки та якістю отриманих результатів. Автоматизація процесу реферування дозволяє значно скоротити час на опрацювання матеріалів та зменшити навантаження на редакторів.

.Додатково варто зазначити, що в умовах інформаційної війни та зростання кількості фейкових новин, автоматизація процесу реферування також може допомогти у швидкій ідентифікації та категоризації потенційно недостовірної інформації, оскільки метадані часто містять важливі маркери для такого аналізу.



3

ПІДХОДИ ДО РЕФЕРУВАННЯ

Extractive

- Передбачає пряме виділення та копіювання найбільш значущих фрагментів вихідного тексту без суттєвих змін їхньої структури та лексики

Abstractive

- Методи дозволяють не просто копіювати фрагменти тексту, а генерувати принципово новий текст, що концентрує зміст оригінального повідомлення

Гібридні

- Поєднують елементи екстрактивного та абстрактного реферування



4

\

МАТЕМАТИЧНА МОДЕЛЬ РЕФЕРУВАННЯ

Математична модель реферування представляється багатовимірною функцією перетворення вихідного тексту Т у реферативний текст R:

$$R = F(T, M, C, P)$$

- де Т – вихідний текстовий масив,
- М – метадані тексту,
- С – контекстуальні параметри,
- Р – прагматичні характеристики повідомлення.

Функціональна залежність включає складні алгоритми семантичного аналізу, що описуються системою нелінійних рівнянь:

$$S(T) = \sum(wi * ti), K(T) = \max(S(T)) / \min(S(T))$$

- де S(T) – семантична вага тексту,
- wi – вагові коефіцієнти слів,
- ti – терміни тексту,
- K(T) – коефіцієнт інформативності.



**ПАРАМЕТРИ СЕМАНТИЧНОГО АНАЛІЗУ
НОВИННИХ ТЕКСТІВ**

Параметр	Позначення	Формула розрахунку	Опис
Інформативність	I	$I = \log(N/n)$	Міра унікальності інформації
Семантична щільність	D	$D = (K / L) * 100\%$	Концентрація смислового навантаження
Контекстуальна релевантність	R	$R = \sum(wi * ci)$	Відповідність контексту

МОДИФІКОВАНИЙ МЕТОД

Послідовність трансформаційних операторів для побудови реферату:

1. Попередня обробка тексту: видалення службових конструкцій, нормалізація лексики, токенизація.
2. Семантичний аналіз: виділення ключових термінів, встановлення семантичних зв'язків.
3. Контекстуальна верифікація: перевірка відповідності метаданим та прагматичним параметрам.

Математична модель контекстуальної валідності визначається формулою:

$$V(R) = \int [C(T) * M(T)] dT,$$

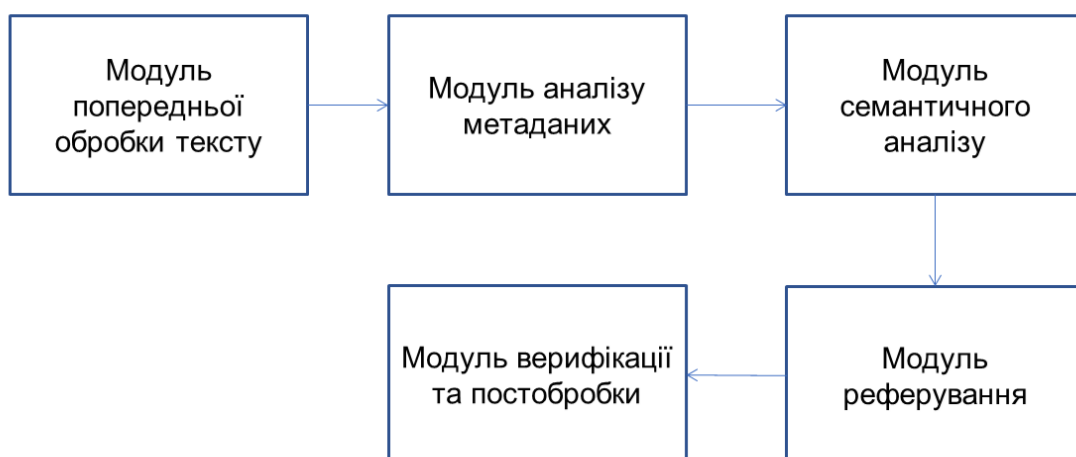
де $V(R)$ – валідність реферату,

$C(T)$ – контекстуальна функція,

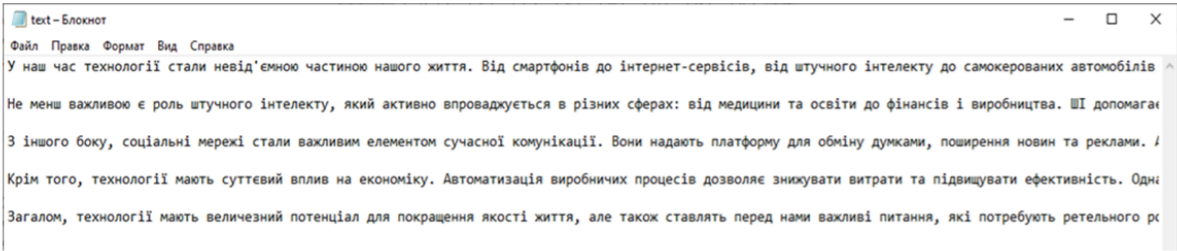
$M(T)$ – функція метаданих.

7

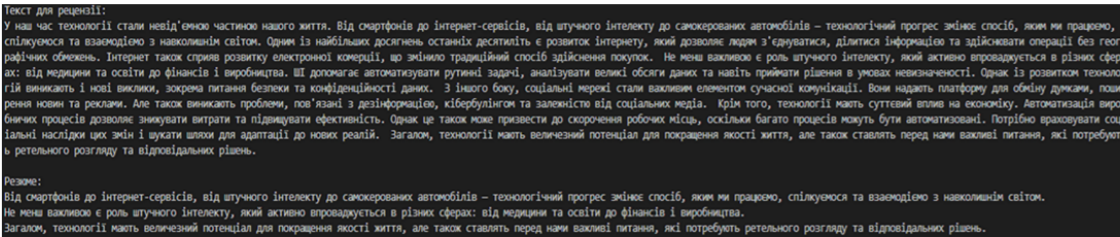
Структура системи автоматизованого реферування



ПРАКТИЧНИЙ РЕЗУЛЬТАТ



Фрагмент вхідного тексту



Результат автоматизованої обробки та формування реферату



ПОРІВНЯННЯ ЧАСУ СТВОРЕННЯ РЕФЕРАТИВ ВРУЧНУ ТА ЗА ДОПОМОГОЮ МЕТОДИКИ

Показник	Ручне створення	Автоматизоване створення	Економія часу
Середній час на один текст	30 хвилин	5 хвилин	83%
Максимальний час на один текст	50 хвилин	7 хвилин	86%
Обсяг текстів за зміну	10 статей	50 статей	500%

ПОРІВНЯННЯ ЯКОСТІ РЕФЕРАТИВ

Критерій	Ручне створення	Автоматизоване створення	Особливості підходу
Точність передачі інформації	95%	92%	Легка перевага у ручній роботі
Відповідність стилю	90%	88%	Потребує адаптації до аудиторії
Структурованість тексту	85%	93%	Перевага у використанні методики

ВИСНОВКИ

1. Проведено аналіз існуючих методів та технологій, що використовуються у реферуванні текстів, зокрема новинних статей, проведено порівняльний аналіз існуючих підходів до автоматизованого реферування.
2. Визначено етапи попередньої обробки новинних текстів, які враховують потреби задачі автоматизованого реферування: очищення від зайвих елементів, таких як HTML-теги, спеціальних символів або рекламних вставок, розбиття на речення та токенизація.
3. Розроблена методика автоматизованого реферування новинних статей на основі метаданих. З урахуванням особливостей текстового контенту враховуються наступні метадані: авторство, час публікації, тематика, ключові слова. Метадані використовуються для точнішого аналізу тексту та врахування його контексту. Для визначення ключових смислових одиниць в роботі використовуються моделі природної обробки мови (NLP), такі як BERT або GPT.
4. Досліджено можливості інтеграції методики у редакційні процеси та адаптації до різних форматів новин. Однією з ключових переваг розробленої методики є можливість її застосування до різних тематичних напрямків і типів текстів, що робить її зручним інструментом для редакцій різного масштабу. Під час реалізації прототипу було підтверджено ефективність використання метаданих для виділення ключових ідей та створення лаконічних рефератів. Це дозволяє значно скоротити час на обробку інформації та зберегти основний зміст повідомлення.

ПУБЛІКАЦІЇ ТА АПРОБАЦІЯ РОБОТИ

Тези доповідей:

1. Гончарова А.А, Золотухіна О.А., Огляд інформаційно-аналітичної діяльності в контексті покращення процесу пошуку інформації на веб-ресурсах та методів реферування статей за допомогою метаданих // Всеукраїнська науково-технічна конференція «Застосування програмного забезпечення в ІКТ». – Київ: ДУІКТ, 2024. – С.166-169.
2. Гончарова А.А, Золотухіна О.А., Аналіз методів реферування новинних статей на основі метаданих за допомогою штучного інтелекту // XXIV Міжнародна науково-практична конференція «Інформаційні технології та безпека» (ІТС-2024). – Київ: НАН України, 2024. – *подано до друку*.

ДОДАТОК Б. ЛІСТИНГИ ПРОГРАМНИХ МОДУЛІВ

```
# main.py
from readers import FileTextReader
from tokenizers import RegexTokenizer
from analyzers import WordFrequencyAnalyzer
from scorers import SimpleSentenceScorer
from summarizer import TextSummarizer

def main():
    """
    Головна функція програми, що ініціалізує компоненти та виконує процес реферування.
    """
    # Шлях до файлу з текстом та кількість речень у резюме
    file_path = 'text.txt'
    num_sentences = 3

    # Ініціалізація компонентів
    reader = FileTextReader()
    tokenizer = RegexTokenizer()
    analyzer = WordFrequencyAnalyzer()
    scorer = SimpleSentenceScorer()
    summarizer = TextSummarizer(reader, tokenizer, analyzer, scorer)

    # Читання та вивід оригінального тексту
    original_text = reader.read(file_path)
    if original_text:
        print(f"Текст для рецензії:\n{original_text}\n")
    else:
        print("Не вдалося прочитати текст для рецензії.\n")
        return

    # Генерація та вивід резюме
    summary = summarizer.summarize(file_path, num_sentences)
    if summary:
        print("Резюме:")
        for sentence in summary:
            print(sentence)
    else:
        print("Не вдалося створити резюме.")

if __name__ == "__main__":
    main()

# analyzers.py
from collections import Counter
from typing import List
from interfaces import AnalyzerInterface

class WordFrequencyAnalyzer(AnalyzerInterface):
    """
```

Клас для аналізу частоти слів у тексті.

"""

```
def analyze(self, words: List[str]) -> Counter:
```

"""

Підраховує частоту кожного слова у списку.

:param words: Список слів.

:return: Counter з частотою кожного слова.

"""

```
word_freq = Counter(words)
```

```
return word_freq
```

interfaces.py

from abc import ABC, abstractmethod

from typing import List

from collections import Counter

```
class TextReaderInterface(ABC):
```

"""

Абстрактний клас, що визначає інтерфейс для читання тексту.

"""

@abstractmethod

```
def read(self, file_path: str) -> str:
```

"""

Читання тексту з файлу.

:param file_path: Шлях до файлу з текстом.

:return: Вміст файлу у вигляді рядка.

"""

```
pass
```

```
class TokenizerInterface(ABC):
```

"""

Абстрактний клас, що визначає інтерфейс для токенизації тексту.

"""

@abstractmethod

```
def tokenize_sentences(self, text: str) -> List[str]:
```

"""

Розділення тексту на речення.

:param text: Вхідний текст.

:return: Список речень.

"""

```
pass
```

@abstractmethod

```
def tokenize_words(self, text: str) -> List[str]:
```

"""

Розділення тексту на слова.

```

        :param text: Вхідний текст.
        :return: Список слів.
        """
        pass

class AnalyzerInterface(ABC):
    """
    Абстрактний клас, що визначає інтерфейс для аналізу тексту.
    """

    @abstractmethod
    def analyze(self, words: List[str]) -> Counter:
        """
        Аналіз частоти слів у списку слів.

        :param words: Список слів.
        :return: Counter з частотою кожного слова.
        """
        pass

class ScorerInterface(ABC):
    """
    Абстрактний клас, що визначає інтерфейс для оцінки речень.
    """

    @abstractmethod
    def score_sentences(self, sentences: List[str], word_freq: Counter) -> List[tuple]:
        """
        Оцінка важливості кожного речення на основі частоти слів.

        :param sentences: Список речень.
        :param word_freq: Counter з частотою слів.
        :return: Список кортежів (речення, оцінка).
        """
        pass

class SummarizerInterface(ABC):
    """
    Абстрактний клас, що визначає інтерфейс для реферування тексту.
    """

    @abstractmethod
    def summarize(self, file_path: str, num_sentences: int) -> List[str]:
        """
        Створення резюме з заданої кількості речень.

        :param file_path: Шлях до файлу з текстом.
        :param num_sentences: Кількість речень у резюме.
        :return: Список речень, що входять до резюме.
        """
        pass

```