

Вінницький національний технічний університет
 Факультет інтелектуальних інформаційних технологій та автоматизації
 Кафедра автоматизації та інтелектуальних інформаційних технологій

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему:

**«Метод автоматизованого генерування альтернативного тексту для
 зображень у вебресурсах»**

Виконав: студент 2 курсу, групи 1АКІТР-24м
спеціальності 174 – Автоматизація,
комп'ютерно-інтегровані технології та
робототехніка

(шифр і назва спеціальності)
Дар'я КАВУНЕЦЬ
 (ПІБ студента)

Керівник: к.т.н., доцент кафедри АІТ

Ярослав КУЛИК
 (науковий ступінь, вчене звання / посада, ПІБ керівника)
 « 11 » 12 2025 р.

Консультант: к.т.н., проф. кафедри АІТ

Ілона БОГАЧ
 Опонент: д.т.н., проф. каф. КН

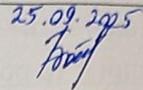
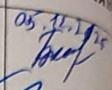
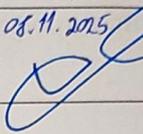
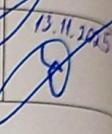
Ярослав ІВАНЧУК
 (науковий ступінь, вчене звання / посада, ПІБ опонента)
 « 12 » 12 2025 р.

Допущено до захисту
 Завідувач кафедри АІТ
д.т.н., проф. Олег БІСІКАЛО
 (науковий ступінь, вчене звання)

« 12 » 12 2025 р.

Вінниця ВНТУ – 2025 рік

6. Консультанти розділів роботи

Розділ змістової частини роботи	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1 – 3	Ілона БОГАЧ к.т.н., професор каф. АІТ	25.09.2025 	05.11.2025 
4	Володимир КОЗЛОВСЬКИЙ к.е.н., доцент каф. ЕПтаВМ	08.11.2025 	13.11.2025 

7. Дата видачі завдання: «25» вересня 2025 року.

КАЛЕНДАРНИЙ ПЛАН

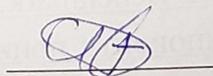
№ з/п	Назва етапів магістерської кваліфікаційної роботи	Строк виконання етапів роботи	Примітки
1	Аналіз предметної області та вимог вебдоступності	25.09 – 05.10.2025	Вик.
2	Розробка алгоритмів автоматичного генерування ALT-текстів	05.10 – 25.10.2025	Вик.
3	Розробка ПЗ для генерації ALT-описів на основі моделі BLIP-2	25.10 – 10.11.2025	Вик.
4	Тестування ПЗ на вибірці зображень та оцінювання якості згенерованих описів	05.11 – 20.11.2025	Вик.
5	Підготовка економічної частини	до 01.12.2025	Вик.
6	Оформлення пояснювальної записки, графічного матеріалу і презентації	20.11 – 03.12.2025	Вик.
7	Попередній захист роботи	до 12.12.2025	Вик.
8	Захист роботи	до 19.12.2025	Вик.

Студент


КАВУНЕЦЬ Д. А.

(прізвище та ініціали)

Керівник роботи


КУЛИК Я. А.

ЗМІСТ

ВСТУП.....	9
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА АКТУАЛЬНОСТІ ТЕМИ	11
1.1 Сучасний стан цифрової доступності: глобальний та український контекст.....	11
1.2 Нормативно-правові вимоги до забезпечення доступності вебресурсів	15
1.2.1 Стандарт WCAG 2.1 і його значення	16
1.2.2 ДСТУ EN 301 549:2022 — український державний стандарт доступності.....	17
1.2.3 Європейський акт про доступність (ЕАА)	18
1.2.4 ADA та Section 508 — американський досвід.....	18
1.2.5 Порівняльний аналіз вимог міжнародних нормативів	19
1.2.6 Значення нормативів для розробників і установ	20
1.3 Проблеми та типові помилки у використанні ALT-текстів на сучасних вебресурсах.....	21
1.4 Роль і функції ALT-текстів у забезпеченні вебдоступності та якісного користувацького досвіду	25
1.4.1 ALT-текст як інструмент інклюзії.....	25
1.4.2 Сприйняття зображень через скрінрідери	26
1.4.3 Типи ALT-текстів та їх функції	26
1.4.4 ALT-текст як елемент UX (користувацького досвіду).....	27
1.4.5 Приклади правильних та неправильних ALT-текстів	28
1.4.6 ALT-тексти в освітньому та науковому середовищі	29
1.5 Методи формування ALT-текстів та необхідність автоматизації	29
1.5.1 Методи формування ALT-текстів: ручні, напівавтоматичні та автоматизовані підходи	29
1.5.2 Основні проблеми автоматизованої генерації ALT-текстів	32
1.5.3 Необхідність автоматизації ALT-текстів в Україні.....	33

1.6 Виклики застосування штучного інтелекту при автоматизованій генерації ALT-текстів	33
1.7 Порівняння існуючих мультимодальних моделей	37
1.8 Висновки до розділу	38
РОЗДІЛ 2 МЕТОДИ ТА ЗАСОБИ АВТОМАТИЗОВАНОГО ГЕНЕРУВАННЯ АЛЬТЕРНАТИВНОГО ТЕКСТУ ДЛЯ ВЕБРЕСУРСІВ.....	41
2.1 Аналіз технологічних особливостей забезпечення доступності зображень та постановка задачі.....	41
2.1.1 Роль зображень та проблеми їх доступності у сучасному вебсередовищі.....	41
2.1.2 Аналіз інструментарію CMS (Joomla, WordPress, Moodle)	42
2.1.3 Постановка задачі дослідження	43
2.2 Розробка методу автоматизованого формування опису зображень	44
2.2.1 Математична формалізація задачі	44
2.2.2 Вибір та обґрунтування використання моделі BLIP2	45
2.2.3 Алгоритм нормалізації тексту.....	45
2.3 Архітектура системи AutoAlt	46
2.3.1 Структурна схема системи	46
2.3.2 Взаємодія компонентів	47
2.4 Програмна реалізація та оптимізація продуктивності	48
2.4.1 Інтеграція з CMS Joomla.....	48
2.4.2 Механізми оптимізації.....	48
2.4.3 Заходи безпеки.....	49
2.4 Висновки до розділу	49
РОЗДІЛ 3 РОЗРОБКА ТА РЕАЛІЗАЦІЯ МОДУЛЯ АВТОМАТИЗОВАНОГО ГЕНЕРУВАННЯ АЛЬТЕРНАТИВНОГО ТЕКСТУ	51
3.1 Концептуальні засади побудови модуля	51
3.2 Архітектурне рішення програмного модуля.....	53

3.2.1 Рівень інтеграції з CMS Joomla.....	54
3.2.2 Рівень серверної логіки (API-сервіс).....	54
3.2.3 Рівень збереження даних: кеш (SQLite).....	55
3.3 Алгоритм функціонування модуля	56
3.3.1 Формалізований опис алгоритму.....	56
3.3.2 Покроковий процес роботи системи	57
3.3.3 Псевдокод алгоритму.....	58
3.3.4 Нормалізація ALT-текстів	59
3.3.5 Відповідність WCAG 2.1	59
3.4 Інтеграція з системами управління контентом.....	61
3.4.1 Механізм роботи з подіями Joomla	61
3.4.2 Аналіз HTML-контенту	61
3.4.3 Взаємодія з API-сервісом	62
3.4.4 Послідовність роботи.....	62
3.4.5 Налаштування плагіну	63
3.4.6 Приклад автоматичного переписування HTML.....	64
3.4.7 Переваги інтеграції	65
3.5 Порівняльний аналіз моделей та ефективності модуля.....	66
3.6 Висновки до розділу	73
РОЗДІЛ 4 ЕКОНОМІЧНИЙ РОЗДІЛ	75
4.1 Технологічний аудит розробленого методу автоматизованого генерування альтернативного тексту для зображень у вебресурсах.....	75
4.2 Розрахунок витрат на розроблення методу автоматизованого генерування alt-тексту	78
4.2.1 Основна заробітна плата.....	78
4.2.2 Додаткова заробітна плата	79
4.2.3 Нарахування на заробітну плату.....	80
4.2.4 Амортизація основних засобів.....	80
4.2.5 Витрати на матеріали	81
4.2.6 Витрати на комплектуючі	81

4.2.7 Витрати на електроенергію	82
4.2.8 Інші витрати.....	82
4.2.9 Сума всіх попередніх статей витрат.....	83
4.2.10 Загальні витрати на розробку методу автоматизованого генерування зображень	83
4.3 Теоретичний розрахунок економічного ефекту від можливої комерціалізації методу	84
4.4 Висновки до розділу	87
ВИСНОВКИ	88
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	90
ДОДАТКИ.....	94
Додаток А (обов'язковий) Технічне завдання Error! Bookmark not defined.	
Додаток Б (обов'язковий) ІЛЮСТРАТИВНА ЧАСТИНА.....	100
Додаток В Лістинг основних функцій програмного забезпечення	107
Додаток Г Акт впровадження.....	111
Додаток Д (обов'язковий) ПРОТОКОЛ ПЕРЕВІРКИ КВАЛІФІКАЦІЙНОЇ РОБОТИ	Error! Bookmark not defined.

ВСТУП

Актуальність роботи. Сьогодні значна частина державних послуг, навчальних матеріалів, новин та соціально важливої інформації надається через вебресурси. Тому питання вебдоступності є не лише технічним, а й суспільно значущим. Для користувачів із порушеннями зору доступ до візуального контенту забезпечується за допомогою альтернативного тексту (ALT-тексту), який є обов'язковою вимогою міжнародного стандарту WCAG 2.1 [1] та національного стандарту ДСТУ EN 301 549:2022 [2].

Попри актуальність цього питання, рівень доступності українських вебресурсів залишається низьким. За результатами моніторингу ПРООН (2023), значна частина державних сайтів не містить коректно заповнених ALT-текстів або не використовує їх зовсім [3]. Причиною цього є необхідність ручного створення текстових описів, що є трудомістким процесом у разі великих обсягів зображень.

Розвиток мультимодальних моделей штучного інтелекту відкриває можливість автоматизації створення ALT-текстів. Це дає змогу підвищити вебдоступність і зменшити навантаження на контент-редакторів, забезпечуючи водночас відповідність вимогам доступності. Тому актуальним є завдання розробити метод та програмний модуль, який зможе автоматично генерувати ALT-тексти українською мовою та працювати локально без передачі даних зовнішнім сервісам, що є критично важливим для державних установ.

Метою роботи є підвищення рівня вебдоступності шляхом розробки методу автоматизованого генерування ALT-текстів на основі мультимодальних моделей штучного інтелекту, оскільки понад 55% вебзображень не мають коректних текстових описів.

Для досягнення мети необхідно **виконати такі задачі:**

1. Проаналізувати вимоги вебдоступності, зокрема WCAG 2.1 [1] та ДСТУ EN 301 549:2022 [2].

2. Дослідити сучасні мультимодальні моделі штучного інтелекту та їх оцінити придатність до автоматизованого опису зображень.

3. Розробити метод генерування ALT-текстів та описати архітектуру програмного модуля.

4. Продемонструвати принцип його інтеграції у веб-сайт та можливість локальної роботи без зовнішніх серверів.

5. Обґрунтувати переваги автоматизації ALT-описів у контексті безбар'єрного цифрового середовища.

Об'єктом дослідження є процес створення альтернативного тексту для зображень у вебресурсах.

Предметом дослідження є методи та моделі автоматизованого генерування ALT-текстів на основі мультимодальних нейронних мереж.

Методи дослідження. У роботі використовуються аналіз нормативних документів, порівняння AI-моделей, методи обробки зображень та природної мови, системний аналіз та моделювання.

Науково-технічний результат полягає у запропонованому підході до локального автоматизованого генерування ALT-текстів на основі мультимодальної моделі VLIIP2, що забезпечує автономну обробку даних і формування україномовних описів відповідно до вимог вебдоступності, що відрізняє його від існуючих онлайн-сервісів.

Практичне значення На основі методу створено програмного забезпечення, яке автоматизує наповнення ALT-текстів на державних, освітніх або корпоративних сайтах, що дозволить зменшити навантаження на контент-редакторів і допоможе виконати вимоги державної політики безбар'єрності.

Апробація результатів роботи. Результати роботи представлено на міжнародній науково-практичній Інтернет-конференції студентів, аспірантів та молодих науковців «Молодь в науці: дослідження, проблеми, перспективи (МН-2025/26)» [43] та впроваджуються в розробки Вінницької обласної військової адміністрації (Додаток Г).

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА АКТУАЛЬНОСТІ ТЕМИ

1.1 Сучасний стан цифрової доступності: глобальний та український контекст

Цифрова трансформація суспільства зумовила стрімке зростання значення вебресурсів у повсякденному житті людей. Сайти та онлайн-платформи сьогодні виконують функції державних сервісів, освітніх інструментів, каналів комунікацій, інформаційних ресурсів та середовищ для соціальної взаємодії. Водночас доступ до цих ресурсів нерівномірно розподілений серед різних груп населення, зокрема серед людей із порушеннями зору. Саме тому питання цифрової доступності вебресурсів набуває стратегічного значення на рівні держав, міжнародних організацій та установ освіти [1].

За даними Всесвітньої організації охорони здоров'я (WHO), у світі налічується понад 2,2 мільярда людей із порушеннями зору, з яких близько 39 мільйонів — повністю незрячі [2]. Кожна з цих людей стикається з бар'єрами, які ускладнюють або унеможливають використання цифрових сервісів. У багатьох випадках недоступність вебресурсів призводить до того, що людина не може читати новини, користуватися державними послугами, реєструватися на прийом до лікаря, навчатися онлайн або здійснювати базові дії, які для зрячих користувачів здаються очевидними.

В Україні, за даними Міністерства соціальної політики, понад 250 тисяч осіб мають значні порушення зору, більшість з яких користуються допоміжними засобами доступу: програмами озвучення (NVDA, JAWS, VoiceOver), брайлівськими дисплеями, збільшувачами та спеціалізованими мобільними застосунками [3]. Для таких користувачів вебсторінка без коректної структури та текстових альтернатив фактично стає недоступною.

У міжнародному контексті проблема доступності набуває все більшої актуальності через розширення електронних сервісів, цифровізацію

адміністративних процесів, академічних платформ та комерційних інструментів. У 2020–2023 роках Європейська комісія та WAI (Web Accessibility Initiative) відзначали, що рівень цифрової доступності у світі залишається критично низьким, попри наявність стандартів і рекомендацій [4]. Ситуація особливо загострюється у країнах, де інформаційні системи активно оновлюються або створюються в умовах швидкої цифрової модернізації.

Одним із наймасштабніших досліджень рівня доступності вебресурсів є щорічний звіт WebAIM Million Report, в якому аналізується 1 000 000 найпопулярніших сайтів світу. Згідно з дослідженням 2023 року, було виявлено, що:

- 96.3% сайтів мають критичні помилки доступності,
- у середньому на кожній домашній сторінці знайдено понад 50 помилок,
- 55% зображень не мають коректного ALT,
- 30% зображень не мають ALT-тексту взагалі,
- багато сайтів містять ALT-тексти з помилками: дублюванням, припущеннями, зайвим описом або порожніми значеннями [5].

Усе це свідчить, що проблема носить системний характер і потребує комплексного підходу.

Недоступність вебресурсів не є лише технічною проблемою — вона безпосередньо впливає на якість життя людей із порушеннями зору. У випадку відсутності альтернативних текстових описів зображень (ALT-text), користувач залишається позбавленим інформації, що критично для:

- освітніх платформ (наприклад, описів схем, графіків, фото лабораторних робіт),
- державних порталів (“Дія”, електронні черги, документи),
- новинних ресурсів,
- бібліотечних каталогів,
- наукових архівів,
- онлайн-магазинів та банківських систем.

Це створює ситуацію цифрової нерівності, коли доступ до інформації залежить не від здібностей людини, а від того, наскільки правильно розробники виконали вимоги доступності.

Стандарти доступності, такі як WCAG 2.1, прямо визначають ALT-текст як обов'язковий елемент структурованого контенту. Зокрема, критерій 1.1.1 “Text Alternatives” говорить про необхідність надавати текстову альтернативу для будь-якого нетекстового контенту, яка забезпечує еквівалентний інформаційний рівень для людей, що не бачать зображення [6]. Відсутність ALT-тексту або його некоректність призводить до незрозуміння структури сторінки, втрати контексту та унеможливлення роботи зі скрінрідером.

Окремо варто зазначити, що в умовах цифровізації державного сектору України вимоги до доступності стають не рекомендаціями, а юридичними зобов'язаннями. Згідно з ДСТУ EN 301 549:2022, ALT-тексти обов'язкові для всіх зображень на державних вебресурсах, окрім декоративних, які мають позначатися пустим ALT. У межах впровадження Європейського акту про доступність (ЕАА), ALT-тексти також набувають ключової ролі в оцінці доступності цифрових сервісів для ринку ЄС [7].

Проблема набуває нового рівня складності з огляду на швидкий розвиток штучного інтелекту та мультимодальних моделей. Попри значний прогрес у розпізнаванні зображень, сучасні моделі (BLIP2, PaLI-X, LLaVA тощо) часто створюють саме описові caption-тексти, а не семантично правильні ALT-тексти. Крім того, вони мають схильність до “галюцинацій” — вигадкування об'єктів, яких на зображенні немає, що унеможлиблює їх безпосереднє використання у державних або освітніх системах без додаткового контролю [8].

У цьому контексті актуальність створення локальної системи автоматичної генерації ALT-текстів — AutoAlt — стає очевидною. Така система має забезпечувати автоматичне формування текстових альтернатив відповідно до WCAG 2.1, працювати локально без зовнішніх API (що важливо для безпеки державних даних), а також інтегруватися в існуючі системи

керування контентом, такі як Joomla. Це дозволить суттєво підвищити рівень доступності українських вебресурсів і сприяти цифровій інклюзії громадян.

Для узагальнення наведених проблем розглянемо результати моніторингу ПРООН (2023-25 рр.), що охоплював понад 200 офіційних сайтів органів влади. Отримані дані демонструють системні порушення вимог WCAG 2.1 та підтверджують критичну потребу у впровадженні інструментів автоматизації перевірки доступності. Узагальнені показники доступності веб-сайтів органів влади України наведено в таблиці 1.1.

Таблиця 1.1 - Основні показники доступності веб-сайтів органів влади України (за даними ПРООН, 2023-25 рр.)

№	Показник оцінки доступності	Частка сайтів, де показник дотримано	Основні проблеми
1	Наявність ALT-текстів для зображень	12 %	ALT відсутній або дублює назву файлу (наприклад, "IMG001.jpg")
2	Коректна навігація з клавіатури	24 %	Не всі елементи активуються через клавішу Tab
3	Контрастність тексту і фону	38 %	Низький контраст у банерах і кнопках
4	Текстові альтернативи для іконок	10 %	Відсутній опис функції іконки
5	Підтримка скрінрідерів	29 %	Відсутні ARIA-мітки, нечитабельні структури таблиць
6	Наявність мовної версії «uk»	65 %	Часто не позначена атрибутом lang="uk"
7	Відповідність WCAG 2.1 рівня А	21 %	Порушення базових вимог доступності
8	Відповідність WCAG 2.1 рівня АА	9 %	Відсутність адаптивних рішень для всіх груп користувачів

Наведені дані підтверджують, що навіть базові вимоги доступності — ALT-тексти, коректна навігація клавіатурою, ARIA-мітки та контрастність — дотримуються менш ніж у третини перевірених сайтів. Це вказує на відсутність системного підходу до доступності та неможливість забезпечити інклюзивність без впровадження стандартизованих інструментів і автоматизованих рішень. У поєднанні з високими вимогами ДСТУ EN 301 549:2022 ці результати підкреслюють актуальність створення локального сервісу AutoAlt.

1.2 Нормативно-правові вимоги до забезпечення доступності вебресурсів

Питання доступності цифрових платформ і вебресурсів регулюється широким спектром міжнародних, європейських та національних нормативних документів. Вони встановлюють стандарти, які гарантують, що кожна людина — незалежно від фізичних можливостей — може користуватися інформаційними технологіями без бар'єрів. У сучасному інформаційному суспільстві такі вимоги перестають бути рекомендаціями та набувають статусу правових зобов'язань, дотримання яких є необхідним як для державних установ, так і для комерційних організацій [9].

У центрі системи цифрової доступності перебуває стандарт WCAG 2.1 (Web Content Accessibility Guidelines), який визначає чіткі правила для створення доступного вебконтенту. Він є основою для національних стандартів ЄС, США та України, а також використовується у більш ніж 120 країнах світу. WCAG визначає 13 принципів і критеріїв доступності, організованих за чотирма фундаментальними принципами POUR: Perceivable (сприйманість), Operable (керованість), Understandable (зрозумілість) та Robust (надійність). Особливе місце у цих стандартах займає вимога щодо текстових

альтернатив для зображень, оскільки цей елемент є ключовим для користувачів із порушеннями зору [10].

1.2.1 Стандарт WCAG 2.1 і його значення

WCAG 2.1 визначає подробиці цифрової доступності на рівнях А, АА та ААА. Переважна більшість державних платформ та великих компаній мають відповідати щонайменше рівню АА, тоді як рівень ААА рекомендований, але не є обов'язковим через надвисокі вимоги.

Серед усіх критеріїв доступності найбільш важливим для користувачів зі слабким зором є критерій 1.1.1 “Text Alternatives”. Він вимагає, щоб кожне зображення мало коректний текстовий опис. Для державних платформ та освітніх установ вимога є безумовною — відсутність ALT-текстів класифікується як порушення стандартів доступності.

Основні правила WCAG щодо ALT-текстів:

- ALT-текст має передавати інформаційний зміст, а не художній опис.
- Заборонено додавати припущення щодо зображення.
- Заборонено використовувати слова “фото”, “зображення”, “малюнок” — це дублює функцію тегу .

- ALT-текст має бути коротким, але інформативним.
- Декоративні зображення повинні мати порожній ALT: alt="".
- Якщо зображення є інформативним (наприклад, графік), ALT має коротко описувати його зміст, а детальний опис повинен бути у тексті поруч.

Згідно з дослідженнями WebAIM, більшість помилок ALT-текстів пов'язані з:

- надмірним описом;
- описом не суті, а деталей;
- припущеннями про емоції, вік, статус людини;
- дубляжем інформації, присутньої в заголовках.

Це відображає ключову проблему: навіть за наявності ALT-тексту він нерідко написаний неправильно й не відповідає вимогам WCAG.

1.2.2 ДСТУ EN 301 549:2022 — український державний стандарт доступності

У 2022 році Україна офіційно запровадила ДСТУ EN 301 549:2022, який є повною імплементацією європейського стандарту EN 301 549. Він встановлює вимоги до доступності:

- веб-сайтів органів влади;
- освітніх платформ;
- інформаційних систем університетів;
- мобільних застосунків;
- державних онлайн-сервісів.

Усі вимоги WCAG 2.1 — включно з критерієм 1.1.1 — є обов'язковими. Це означає, що відсутність ALT-текстів на сайті державної установи або університету є порушенням стандартів і може бути підставою для адміністративних наслідків.

Ключові вимоги стандарту стосуються:

- структури вебсторінок;
- семантичної розмітки;
- доступності мультимедійних матеріалів;
- текстових альтернатив;
- навігації;
- контрастності;
- масштабування;
- доступності для скрінрідерів.

У стандарті спеціально зазначено, що відсутність ALT-текстів блокує доступність контенту для осіб з порушеннями зору, а тому має бути виправлена на етапі розробки або оновлення вебресурсу [11].

1.2.3 Європейський акт про доступність (ЕАА)

European Accessibility Act (Directive 2019/882/EU) — один із найважливіших документів ЄС у сфері цифрової доступності. Він набуває чинності в повному обсязі у 2025 році, а це означає, що всі цифрові продукти, які працюють на території ЄС, повинні відповідати його вимогам.

До таких продуктів належать:

- освітні та інформаційні платформи;
- онлайн-магазини;
- електронні банкінги;
- термінали самообслуговування;
- мобільні застосунки;
- державні та приватні веб-сайти.

ALT-тексти є одним із ключових критеріїв ЕАА.

Порушення вимог може призвести до:

- штрафів;
- обмеження доступу на європейський ринок;
- зобов'язання провести повний аудит доступності;
- для університетів — відмови у міжнародних грантах та програмах.

1.2.4 ADA та Section 508 — американський досвід

У США регулюванням доступності займаються два основні закони:

- ADA (Americans with Disabilities Act)

- Section 508 Rehabilitation Act

Обидва зобов'язують державні установи, університети й компанії забезпечувати відповідність WCAG.

Згідно з офіційними звітами Міністерства юстиції США:

- у 2022–2023 роках було подано понад 4 000 судових позовів через недоступність веб-сайтів;

- відсутність або неправильність ALT-текстів входить у ТОП-3 найпоширеніших порушень;

- середній розмір штрафів і компенсацій — від 20 000 до 50 000 доларів [12].

Цей досвід демонструє, що ALT-тексти — не дрібниця, а юридично значима складова доступності.

1.2.5 Порівняльний аналіз вимог міжнародних нормативів

Для систематизації вимог до текстового опису зображень було проаналізовано ключові міжнародні та національні стандарти. Результати порівняльного аналізу щодо обов'язковості наявності ALT-текстів та необхідних рівнів відповідності наведено в таблиці 1.2.

Висновок: Вимога щодо наявності ALT-текстів є універсальною, юридично обов'язковою та присутня у всіх розглянутих міжнародних стандартах доступності.

Таблиця 1.2 — Порівняльний аналіз вимог міжнародних нормативів щодо наявності альтернативного тексту (ALT)

Нормативний документ	Обов'язковість ALT	Рівень відповідності	Сфера застосування
WCAG 2.1	Так	A / AA / AAA	Усі вебресурси
EN 301 549	Так	AA	ЄС
ДСТУ EN 301 549	Так	AA	Україна
EAA 2019/882	Так	AA	ЄС (державний та приватний сектори)
ADA	Так	WCAG AA	США
Section 508	Так	WCAG AA	США (державний сектор)

1.2.6 Значення нормативів для розробників і установ

Для державних і освітніх платформ виконання нормативів означає:

- можливість участі у міжнародних програмах та грантах;
- відповідність вимогам ЄС;
- захист від юридичних ризиків;
- підвищення якості цифрового середовища;
- збільшення доступності ресурсів для студентів і громадян.

Для розробників це гарантує:

- стандартизовані підходи до створення інтерфейсів;
- можливість інтеграції з міжнародними системами;
- юридичний захист продукту.

1.3 Проблеми та типові помилки у використанні ALT-текстів на сучасних вебресурсах

Попри існування чітких міжнародних стандартів доступності, таких як WCAG 2.1, значна частина вебресурсів містить помилки або зовсім не використовує ALT-тексти. Некоректне застосування текстових альтернатив створює суттєві бар'єри для користувачів зі зниженим зором, які взаємодіють з контентом за допомогою програм екранного озвучення (NVDA, JAWS, VoiceOver). Аналіз причин цих помилок є критично важливим для формування ефективних підходів до автоматизації створення ALT-текстів та розробки системи AutoAlt.

Згідно з WebAIM Million Report 2024, серед 1 000 000 найбільш відвідуваних сайтів у світі:

- 55% зображень не мають ALT-текстів,
- 30% містять порожній або некоректний ALT,
- надмірно довгі або описові ALT-тексти зустрічаються в 12% випадків,
- 6% нечитабельні або містять технічні назви файлів [13].

Ці дані демонструють, що навіть за наявності рекомендацій проблема залишається системною.

Розглянемо основні проблеми та типові помилки у використанні ALT-текстів на сучасних вебресурсах:

1. Відсутність ALT-текстів як одна з найпоширеніших проблем
2. Надмірний опис або “поетизація” ALT-текстів
3. ALT-текст із припущеннями
4. Неповні або дублюючі ALT-тексти
5. Технічні та нерелевантні ALT-тексти
6. Неправильне використання порожнього ALT
7. ALT-текст не відповідає ролі зображення
8. Вплив помилок ALT-текстів на користувачів

9. Узагальнення: чому це проблема для держави, освіти та бізнесу

Розглянемо кожен з них більш детально.

Відсутність ALT-текстів як одна з найпоширеніших проблем.

Найсерйознішим порушенням є повна відсутність ALT-тексту. У такому випадку програма озвучення пропускає зображення або відтворює його ім'я файлу, наприклад:

«IMG_20230101_0034.jpg»

Для незрячого користувача така інформація не несе жодної цінності.

Відсутність ALT-тексту особливо критична на:

- державних порталах;
- освітніх ресурсах;
- інформаційних сайтах;
- інтернет-магазинах;
- новинних платформах;
- сайтах університетів та бібліотек.

Наприклад, якщо на сторінці курсу з'являється графік без опису, студент із порушенням зору фактично позбавлений доступу до навчального матеріалу.

Надмірний опис або "поетизація" ALT-текстів. Частою проблемою є надмірно детальні або художні описи, які не відповідають вимогам WCAG.

Некоректний ALT-приклад:

«Красива усміхнена дівчина в чудовому літньому платті стоїть на фоні заходу сонця...»

Такий текст не лише суб'єктивний, але й надмірно довгий. Скринрідери зачитують ALT повністю, тому перевантаження інформацією уповільнює взаємодію з сайтом.

Правильний варіант:

«Жінка стоїть на вулиці під час заходу сонця».

ALT-текст із припущеннями. ALT-текст не має містити інформації, яку неможливо підтвердити зображенням. Програми озвучення відтворюють те, що написано, а не те, що може бути правдою.

Приклад помилки:

«Студент готується до іспиту і виглядає втомленим».

Тут присутні:

- припущення про статус людини;
- припущення про її емоційний стан;
- неправильне уявлення про контекст.

Правильний варіант:

«Чоловік сидить за ноутбуком у кімнаті».

Неповні або дублюючі ALT-тексти. ALT-тексти часто дублюють заголовок або навколишній текст:

Приклад дубляжу:

Заголовок: “Результати опитування студентів”

ALT: “Результати опитування студентів”

Це помилка — ALT має описувати саме зображення, а не повторювати текст поруч.

Технічні та нерелевантні ALT-тексти. ALT-тексти на кшталт:

- “image123.jpg”

“picture”

- “photo1”

є типовими, але абсолютно неприпустимими.

У таких випадках користувач отримує нуль інформації.

Неправильне використання порожнього ALT. WCAG чітко визначає: декоративні зображення повинні мати alt="", щоб їх пропускали скринрідери.

Помилка:

alt="декоративна лінія"

alt="фонова іконка"

alt="***"

У цих випадках скринрідер зайвий раз відволікає користувача на непотрібні непринципові елементи.

ALT-текст не відповідає ролі зображення. Наприклад, на кнопці “Пошук” використана іконка лупи.

Помилковим буде ALT:

«Лупа»

Правильний варіант:

«Пошук»

Це називається функціональний ALT.

Вплив помилок ALT-текстів на користувачів Помилки ALT-текстів спричиняють:

- втрату частини або повного контенту;
- падіння швидкості сприймання сторінки;
- дезорієнтацію користувача;
- неможливість проходити курси або використовувати сервіси;
- неправильне розуміння матеріалів;
- повну недоступність графіків/таблиць;
- погіршення академічної успішності (в освітніх середовищах).

Скринрідери читають ALT-тексти послідовно, тому кожна зайва помилка множить для сотень студентів, які читають сторінку щодня.

Узагальнення: чому це проблема для держави, освіти та бізнесу.

Некоректні ALT-тексти — це не просто технічна вада. Це:

- юридичне порушення;
- бар’єр для доступу до освіти;
- ризик для державних систем;
- підстава для скарг та аудитів;
- відмова у фінансуванні міжнародних проектів;
- втрата реальних користувачів (до 15% аудиторії).

1.4 Роль і функції ALT-текстів у забезпеченні вебдоступності та якісного користувацького досвіду

ALT-текст (alternative text) є фундаментальним елементом вебдоступності та одним із ключових інструментів комунікації між авторами вебконтенту і користувачами з порушеннями зору. Його значення виходить далеко за межі формального дотримання вимог WCAG 2.1 — ALT-текст відіграє визначальну роль у формуванні користувацького досвіду (UX), доступності навчальних матеріалів, інклюзивності цифрового середовища та загальної якості інформаційної архітектури вебресурсу.

У контексті цифрової інклюзії ALT-текст виступає способом забезпечення еквівалентності доступу до інформації, коли користувач, який не бачить зображення, отримує текстовий опис, що виконує ту саму функцію, що й графічний елемент. Таким чином, ALT-текст є мостом між візуальним контентом та користувачем, який покладається на голосовий синтез або брайлівський дисплей.

1.4.1 ALT-текст як інструмент інклюзії

Особливість ALT-текстів полягає в тому, що вони не лише описують зображення, а забезпечують рівність доступу. Для мільйонів користувачів зі зниженою гостротою зору альтернативний текст — єдина можливість:

- зрозуміти інформацію на сторінці;
- перейти за посиланням, якщо зображення є кнопкою;
- інтерпретувати структуру навчальних матеріалів;
- орієнтуватися у формах, графіках та схемах;
- сприймати мультимедійний контент;
- взаємодіяти з державними сервісами.

Без ALT-текстів користувач не може отримати інформацію, що робить сторінку фактично недоступною, навіть якщо вона повністю працездатна для зрячих відвідувачів.

1.4.2 Сприйняття зображень через скрінрідери

Програми озвучення (NVDA, JAWS, VoiceOver) послідовно читають елементи сторінки, включно з ALT-текстами.

Скорочений або неправильний ALT впливає на сприйняття контенту:

- надмірно детальні ALT перевантажують користувача;
- занадто короткі ALT не дають інформації;
- неправильні ALT вводять користувача в оману;
- відсутні ALT створюють прогалини у змісті.

Особливо важливо це у навчальних матеріалах, де зображення можуть містити критично важливі елементи резюме, моделі, схеми або наукові візуалізації.

1.4.3 Типи ALT-текстів та їх функції

Розробники та контент-мейкери мають розуміти, що ALT-тексти бувають різних типів, залежно від функції зображення:

а) Функціональні ALT-тексти

Призначені для зображень, які мають дію або роль: кнопки, іконки, елементи навігації.

Приклад:

- X → «Меню»
- Іконка лупи → «Пошук»

Важливо: описувати не що намальовано, а що це робить.

b) Інформаційні ALT-тексти

Призначені для фото, ілюстрацій, графіків, що несуть зміст.

Приклади:

- «Студент працює за ноутбуком у бібліотеці»
- «Графік зростання відвідуваності за 2022–2023 роки»

c) Описові ALT-тексти для складних зображень

Схеми, інфографіка, діаграми потребують короткого ALT + детального тексту під зображенням.

Приклад ALT:

- «Діаграма структури фінансування університету»

Детальний опис розміщується нижче як підпис до зображення — це одна з вимог WCAG.

d) Декоративні ALT-тексти (порожні)

Якщо зображення не має інформаційної цінності:

`alt=""`

Це дозволяє скрінрідерам пропустити декоративний елемент і пришвидшити навігацію.

1.4.4 ALT-текст як елемент UX (користувацького досвіду)

Сприйняття сторінки через скрінрідер відрізняється від візуального перегляду:

- користувач читає контент лінійно, а не бачить цілісну структуру;
- кожне слово має значення — тому ALT-фрагменти мають бути точними;
- некоректні ALT можуть “зламати” семантичну структуру сторінки;
- велика кількість декоративних ALT уповільнює роботу сторінки.

Провідні UX-компанії (наприклад, Nielsen Norman Group) показують, що правильно створені ALT-тексти скорочують час взаємодії зі сторінкою на 20–40% та зменшують когнітивне навантаження на користувачів.

1.4.5 Приклади правильних та неправильних ALT-текстів

Для уникнення типових помилок при генерації чи написанні описів важливо розрізняти інформативний та надлишковий контент. У таблиці 1.3 наведено порівняння некоректних та правильних варіантів ALT-текстів для різних типів графічних елементів.

Таблиця 1.3 — Приклади правильного та неправильного оформлення ALT-текстів

Тип зображення	Неправильно (Типові помилки)	Правильно (Рекомендований варіант)
Фото людини	«Гарна дівчина» (суб'єктивно)	«Жінка сидить за ноутбуком у кімнаті»
Графік	«Графік» (занадто загально)	«Графік зростання трафіку за 2021–2023 роки»
Іконка	«Іконка меню» (дублювання слова «іконка»)	«Меню» (опис функції)
Декор	«Лінія» (опис елемента, що не несе змісту)	alt="" (порожній атрибут)
Фото природи	«Краєвид» (неінформативно)	«Гори під час заходу сонця»

1.4.6 ALT-тексти в освітньому та науковому середовищі

У вищих навчальних закладах ALT-тексти виконують ще важливішу роль:

- забезпечують рівний доступ до навчальних матеріалів;
- дозволяють студентам з порушеннями зору працювати з презентаціями та лабораторними роботами;
- забезпечують відповідність університету європейським вимогам освіти;
- впливають на акредитацію освітніх програм.

Погано описані або відсутні ALT-тексти є бар'єром для студентів і можуть впливати на академічну успішність.

1.5 Методи формування ALT-текстів та необхідність автоматизації

1.5.1 Методи формування ALT-текстів: ручні, напівавтоматичні та автоматизовані підходи

ALT-текст є критично важливим елементом вебдоступності, проте створення якісних текстових альтернатив може бути складним завданням навіть для досвідчених контент-менеджерів та розробників. На практиці використовуються три основні підходи формування ALT-текстів: ручний, напівавтоматичний та автоматизований. Кожен з них має свої переваги, недоліки та сфери застосування. Аналіз цих методів є необхідним для обґрунтування вибору технічного рішення, що пропонується у межах даної магістерської роботи.

Ручне формування ALT-текстів. Ручне написання ALT-текстів — це традиційний підхід, що залишається найбільш точним, коли його виконують кваліфіковані спеціалісти з доступності або автор контенту.

Переваги ручного підходу:

- a) Висока точність. Людина може коректно описати контекст і зміст зображення, уникаючи помилок, характерних для нейромереж.
- b) Розуміння ролі зображення. ALT-текст залежить не лише від того, що зображено, а й для чого це зображення. Людина це враховує.
- c) Відсутність “галюцинацій”. Людина не вигадує деталі, яких немає, на відміну від ШІ.

Недоліки ручного підходу:

- a) Дуже висока трудомісткість. Дослідження W3C показують, що створення якісного ALT-тексту займає від 20 до 60 секунд на одне зображення [21].
- b) Низька масштабованість. Університетські сайти, освітні матеріали та новинні портали можуть містити тисячі зображень. Ручне створення ALT стає фізично неможливим.
- c) Людський фактор. Не всі автори контенту знають стандарти WCAG → велика кількість помилок.
- d) Нестабільна якість. ALT-тексти, створені різними людьми, часто суттєво відрізняються за стилем та рівнем деталізації.

Напівавтоматичні підходи. Напівавтоматичні методи використовують алгоритми, які пропонують варіант ALT-тексту, але потребують ручного редагування.

Поширені інструменти:

- автоматичні підказки у CMS;
- використання OCR для текстових зображень;
- бібліотеки з вбудованими caption-модулями;
- прості моделі машинного навчання.

Переваги:

- швидше, ніж ручне написання;
- корисно для великих обсягів медіа;
- дозволяє редактору виправити помилки;

- знижує навантаження на контент-менеджерів.

Недоліки:

- у 60–70% випадків запропонований ALT потребує доопрацювання;
- моделі часто описують те, що бачать, а не функцію зображення;
- можливі помилки розпізнавання;
- не підходить для державних сайтів, де потрібна гарантована коректність.

Автоматизовані підходи (на основі нейромереж). Сучасні мультимодальні моделі (BLIP2, LLaVA, PaLI-X, InstructBLIP, Kosmos-2) здатні генерувати текстові описи зображень. Вони використовуються компаніями Meta, Google, Microsoft для автоматичного captioning. Однак caption \neq ALT-текст.

Основна проблема: Caption описує картинку “поетично” або “емоційно”, але не функціонально. ALT-текст повинен бути коротким, нейтральним, точним і відповідати WCAG.

Порівняння методів формування ALT-текстів наведено в таблиці 1.4.

Таблиця 1.4 - Порівняльний аналіз підходів формування ALT-текстів

Критерій	Ручний підхід	Напівавтоматичний	Автоматизований
Точність	Висока	Середня	Низька–середня
Швидкість	Низька	Середня	Висока
Масштабованість	Низька	Середня	Висока
Ризик помилок	Низький	Середній	Високий
Вимоги до фахівця	Високі	Середні	Низькі
Відповідність WCAG	Висока	Часткова	Нестабільна
Сумісність з держ. стандартами	Повна	Часткова	Не гарантована

Ручний підхід є еталоном якості та відповідності стандартам, але він не підходить для великих обсягів даних через низьку швидкість.

Автоматизований підхід ідеальний для масштабування та швидкості, але потребує суворого контролю якості (верифікації), оскільки не гарантує точності.

Напівавтоматичний підхід виступає компромісом, поєднуючи швидкість інструментів та контроль людини.

1.5.2 Основні проблеми автоматизованої генерації ALT-текстів

Сучасні моделі мають такі характерні недоліки:

а) Галюцинації мультимодальних моделей

Модель може вигадати:

- предмети;
- людей;
- емоції;
- дії;
- кольори.

Це критично для освітніх та державних систем.

б) Неправильна інтерпретація функції зображення

Наприклад, іконку «Меню» модель описує як:

“три горизонтальні лінії”

хоча правильний ALT — “Меню”, бо важлива роль, а не вигляд.

в) Надмірно довгі описи

ШІ генерує:

“A beautiful woman standing in a sunny field...”

Тоді як ALT має бути коротким:

“Жінка стоїть у полі”.

г) Юридичні ризики

Некоректний ALT на державному сайті = порушення ДСТУ EN 301 549
порушення WCAG 2.1 ризик аудиту ЄС скарги користувачів

1.5.3 Необхідність автоматизації ALT-текстів в Україні

Україна активно рухається у напрямку цифровізації та інтеграції до європейського простору. Це означає:

- вимоги доступності стануть жорсткішими;
- кількість контенту постійно зростає;
- ручне створення ALT неможливе на практиці;
- університети, державні сайти та медіа потребують автоматизованих рішень.

Система AutoAlt, яка:

- працює локально,
- генерує короткі WCAG-сумісні ALT-описи,
- не передає дані у хмару,
- інтегрується в CMS Joomla,
- є абсолютно необхідною для відповідності нормам ЄС та WCAG.

1.6 Виклики застосування штучного інтелекту при автоматизованій генерації ALT-текстів

Стрімкий розвиток мультимодальних моделей штучного інтелекту (ШІ), здатних аналізувати зображення та формувати текстові описи, відкриває нові можливості у сфері забезпечення цифрової доступності. Моделі типу BLIP2, LLaVA, PaLI-X, Kosmos-2, InstructBLIP демонструють значні досягнення у задачах візуального розпізнавання та генерування описів. Проте їх застосування у контексті створення саме ALT-текстів часто виявляється

проблемним. ALT-текст — це не просто опис зображення, а структурований, нейтральний, функціональний елемент доступності, суворо регламентований стандартами WCAG 2.1 та ДСТУ EN 301 549.

Тому перед розробниками стоїть завдання: використати можливості ШІ, уникаючи типових помилок, забезпечивши відповідність нормативним вимогам і мінімізувавши ризики помилкових або небезпечних генерацій. Нижче розглядаються ключові виклики, пов’язані із застосуванням ШІ в задачі формування ALT-текстів.

Розглянемо основні проблеми:

1. Проблема галюцинацій нейромереж
2. Проблема невідповідності між caption і ALT-текстом
3. Надмірна деталізація або “літературизація” описів
4. Обмеженість контекстного розуміння
5. Етичні ризики та небажана інтерпретація
6. Проблеми приватності та безпеки даних
7. Проблема продуктивності та апаратних обмежень

Розглянемо їх більш детально.

Проблема галюцинацій нейромереж. Однією з найважливіших проблем мультимодальних моделей є явище так званих галюцинацій — ситуацій, коли модель описує об’єкти, яких фактично немає на зображенні.

Типові приклади галюцинацій:

- вигадування неіснуючих предметів;
- приписування людині певних емоцій або намірів;
- помилкове визначення статі, віку, етнічної належності;
- вигадування тексту, який нібито присутній на зображенні;
- присвоєння зображенню хибного контексту.

Це є критичним порушенням вимог WCAG, оскільки ALT-текст має надавати точну і об’єктивну інформацію, а не припущення або оцінки. Дослідження Google Research та Meta AI підтверджують, що навіть найкращі

моделі можуть допускати галюцинації у 25–40% випадків при генерації caption-текстів [26].

У контексті ALT-тексту помилка може зробити зображення не просто недоступним — а хибно зрозумілим, що руйнує інформаційну структуру сторінки.

Проблема невідповідності між caption і ALT-текстом. Мультиmodalні моделі навчені формувати caption — опис, орієнтований на художню чи контекстну інтерпретацію — а не функціональний ALT-текст.

Caption-текст відповідає на запитання: «Що я бачу?»

ALT-текст відповідає на питання: «Яку інформацію ця картинка несе для користувача?»

Caption може бути довгим, емоційним, суб'єктивним: «Молода усміхнена жінка милується заходом сонця на березі моря».

ALT повинен бути коротким і нейтральним: «Жінка стоїть на березі моря під час заходу сонця».

Це принципово різні речі, тому використання ШІ без спеціальної адаптації приводить до порушення стандартів WCAG 2.1.

Надмірна деталізація або “літературизація” описів. ШІ часто генерує описи, які містять:

- зайві деталі (“красива”, “яскравий”, “затишний”);
- недоречні припущення (“схоже, що вони святкують”);
- емоційні оцінки (“розкішний краєвид”).

ALT-текстами WCAG категорично забороняється.

Скринрідери зачитують текст дослівно — чим довший ALT, тим складніше користувачу сприймати сторінку.

Це погіршує UX і порушує принцип Understandable стандарту WCAG.

Обмеженість контекстного розуміння. ШІ не має повного контексту сторінки.

Навіть найкращі моделі:

- не знають, де розміщене зображення на сторінці;

- не розуміють ролі картинки у контенті;
- не можуть вирішити, чи картинка декоративна;
- не можуть оцінити, чи потрібен короткий ALT, чи детальний опис.

Наприклад, модель однаково опише іконку “лупа” і картинку лупи у підручнику фізики — але перша потребує ALT “Пошук”, а друга — “Лупа”.

Етичні ризики та небажана інтерпретація. ШІ може несвідомо генерувати упереджені або дискримінаційні формулювання, пов’язані з:

- статтю;
- віком;
- етнічністю;
- соціальними стереотипами.

Це суперечить європейським нормам інклюзії та WCAG.

Проблеми приватності та безпеки даних. Більшість готових рішень ШІ для генерації ALT-текстів працюють через хмарні API:

- Google Vision
- AWS Rekognition
- Microsoft Azure Cognitive Services

Для державних установ та університетів України це неприйнятно, тому що:

- дані (включно з фото) передаються на зовнішні сервери;
- це потенційне порушення закону “Про захист персональних даних”;
- Європейський акт про цифрову безпеку вимагає локальних рішень.

AutoAlt вирішує цю проблему, оскільки працює локально, не передає інформацію у хмару і повністю відповідає вимогам безпеки.

Проблема продуктивності та апаратних обмежень. Мультиmodalьні моделі зазвичай потребують:

- GPU,
- багато оперативної пам’яті,
- великих обчислювальних ресурсів.

1.7 Порівняння існуючих мультимодальних моделей

Порівняння продуктивності мультимодальних моделей на CPU наведено в таблиці 1.5.

Таблиця 1.5 – Порівняння продуктивності мультимодальних моделей на CPU

Модель	Середній час генерації на CPU	Вимоги до апаратного забезпечення
BLIP2	2–4 секунди	Середні (придатна для локальних серверів)
LLaVA	20–40 секунд	Дуже високі (ефективна лише на GPU)
PaLI-X	> 1 хвилини	Нереалістичні для локальних систем

Тому для реальних державних платформ вибір моделі повинен базуватися не лише на точності, а й на швидкості, що є важливим аргументом на користь BLIP2 у твоїй системі.

Порівняння сучасних мультимодальних моделей наведено в таблиці 1.6.

Таблиця 1.6 — Порівняння мультимодальних моделей щодо придатності для ALT-текстів

Характеристика	BLIP2	LLaVA	PaLI-X	Kosmos-2
Точність описів	Середня	Висока	Дуже висока	Середня
Ризик галюцинацій	Низький	Середній	Високий	Середній
Швидкість роботи	Висока	Низька	Дуже низька	Середня
Потреби у GPU	Низькі	Високі	Дуже високі	Високі
Сумісність з локальними системами	Висока	Низька	Дуже низька	Середня
Придатність для ALT	Висока	Середня	Низька	Середня

BLIP2 виглядає як найбільш збалансоване рішення для практичного застосування (генерації ALT-текстів у реальному часі). Вона швидка, має низькі вимоги до "заліза" та низький ризик галюцинацій, що критично для доступності. PaLI-X забезпечує найвищу якість, але її "важкість" (потреби в ресурсах та низька швидкість) робить її малопридатною для масової автоматичної обробки зображень на звичайних серверах. LLaVA демонструє високу точність, але має проблеми зі швидкістю та сумісністю, що може бути перешкодою для інтеграції у вебресурси.

Використання ШІ для генерації ALT-текстів має великий потенціал, але супроводжується рядом викликів:

- ризик галюцинацій;
- недостатня відповідність WCAG;
- небезпека надмірних або упереджених описів;
- технічні й апаратні обмеження;
- питання безпеки даних;
- відсутність контекстного розуміння.

Саме ці виклики визначають актуальність розробки спеціалізованих локальних систем, які адаптовані спеціально під вимоги WCAG та характеристики українських державних і освітніх платформ.

1.8 Висновки до розділу

У першому розділі було проведено комплексний аналіз сучасного стану цифрової доступності та ролі ALT-текстів як ключового інструменту забезпечення рівного доступу до інформації для користувачів з порушеннями зору. Встановлено, що доступність вебресурсів є не лише технічним завданням, а важливою соціальною, правовою та етичною вимогою, яка визначає можливості людей із порушеннями зору користуватися цифровими сервісами на рівних умовах.

Аналіз статистичних даних міжнародних організацій, зокрема WHO та WebAIM, показав, що значна частина вебресурсів у світі залишається недоступною. Більше ніж 55% зображень не мають ALT-текстів, а приблизно третина містить некоректні або порожні текстові альтернативи. Це створює суттєві перешкоди для користувачів, які покладаються на скрінрідери, та призводить до високого рівня цифрової нерівності в суспільстві. Ситуація ускладнюється ще й тим, що багато авторів контенту не володіють належними знаннями щодо стандартів доступності, що призводить до широкого спектра типових помилок у створенні ALT-текстів.

Результати аналізу нормативної бази — WCAG 2.1, ДСТУ EN 301 549:2022, European Accessibility Act, ADA та Section 508 — свідчать, що вимоги щодо надання текстових альтернатив є універсальними, обов'язковими та суворо регламентованими. Відсутність або неправильність ALT-текстів є прямим порушенням міжнародних і національних стандартів, що може призводити до юридичної відповідальності, обмежень у фінансуванні, втрати користувачів та погіршення іміджу організації. Таким чином, ALT-текст є не факультативним елементом вебдизайну, а невід'ємною складовою цифрової інклюзії.

У підрозділах 1.3 та 1.4 було розкрито типові помилки у використанні ALT-текстів та їхній вплив на доступність і користувацький досвід. Показано, що некоректні ALT-тексти можуть бути не менш шкідливими, ніж їхня повна відсутність, оскільки неправильний опис призводить до спотворення змісту та порушення логіки сприйняття сторінки. Правильно створений ALT-текст виконує функцію інформаційного та функціонального еквівалента зображення, забезпечує ефективну взаємодію зі скрінрідерами та покращує загальний UX, а також позитивно впливає на SEO-показники вебресурсу.

Окрему увагу було приділено аналізу сучасних підходів до формування ALT-текстів — ручних, напівавтоматичних та автоматизованих. Встановлено, що ручний підхід забезпечує найвищу точність, але є вкрай трудомістким та маломасштабованим. Напівавтоматичні методи зменшують навантаження, але

все ще потребують значної участі людини. Автоматизовані підходи на основі мультимодальних моделей ШІ мають великий потенціал, але наразі супроводжуються низкою проблем: галюцинаціями, недостатнім розумінням контексту, надмірною деталізацією, непередбачуваністю результатів та високими вимогами до обчислювальних ресурсів.

У підрозділі 1.6 було детально досліджено виклики застосування ШІ для генерації ALT-текстів. Зокрема, виявлено, що сучасні мультимодальні моделі формують caption-описи, які не відповідають вимогам WCAG, та демонструють нестабільність у точності, що робить їх непридатними для безпосереднього використання в державних або освітніх вебресурсах без спеціальної адаптації. Також акцентовано увагу на питаннях приватності даних, що є критично важливими для України й ЄС.

Таким чином, результати першого розділу дозволяють зробити важливий висновок: існуючі методи створення ALT-текстів не забезпечують ефективного, масштабованого та нормативно коректного рішення, особливо в умовах великих обсягів контенту та потреби в автоматизації. Ручні й напівавтоматичні підходи не можуть бути застосовані на рівні державних порталів та університетських систем, а сучасні ШІ-моделі потребують адаптації та контролю якості.

Отже, необхідність розроблення спеціалізованого локального сервісу AutoAlt є обґрунтованою і впливає з аналізу нормативної бази, технічних обмежень, соціальних потреб та викликів застосування ШІ. У наступному розділі буде розглянуто технічні вимоги майбутньої системи, архітектуру рішення, особливості інтеграції з CMS Joomla та підходи до забезпечення коректності, швидкості й безпеки автоматичної генерації ALT-текстів.

РОЗДІЛ 2 МЕТОДИ ТА ЗАСОБИ АВТОМАТИЗОВАНОГО ГЕНЕРУВАННЯ АЛЬТЕРНАТИВНОГО ТЕКСТУ ДЛЯ ВЕБ-РЕСУРСІВ

2.1 Аналіз технологічних особливостей забезпечення доступності зображень та постановка задачі

У даному підрозділі здійснюється аналіз технологічного середовища, у якому функціонує розроблювана система AutoAlt. Розглядаються процеси обробки зображень у вебпросторі, специфіка формування текстових альтернатив та обмеження існуючих систем управління контентом (CMS). На основі аналізу формулюється детальна постановка задачі дослідження.

2.1.1 Роль зображень та проблеми їх доступності у сучасному вебсередовищі

Зображення є невід’ємною складовою сучасного вебконтенту. За даними HTTP Archive, графічні елементи становлять у середньому 48–52 % від загального обсягу переданих даних вебсторінки [31]. Коректне оформлення графіки безпосередньо впливає на доступність ресурсу для користувачів з порушеннями зору, які використовують допоміжні технології (скрінрідери NVDA, JAWS, VoiceOver) [32].

Згідно з міжнародним стандартом WCAG 2.1 (критерій 1.1.1), будь-який нетекстовий контент повинен мати текстову альтернативу, яка виконує еквівалентну функцію [10]. Проте аналіз існуючих українських освітніх та державних ресурсів виявив критичні проблеми:

1. Відсутність опису: Близько 60–78 % зображень не мають атрибуту alt або він порожній [5].

2. Технічні описи: Використання назв файлів (наприклад, IMG_2024.jpg) замість змістовного опису, що не несе користі для користувача.

3. Надмірність: Дублювання інформації, яка вже є в контексті сторінки.

2.1.2 Аналіз інструментарію CMS (Joomla, WordPress, Moodle)

Найпоширеніші в державному та освітньому секторі України CMS мають суттєві недоліки в контексті автоматизації доступності. Порівняльний аналіз можливостей популярних систем управління контентом наведено в таблиці 2.1.

Таблиця 2.1 — Порівняння популярних CMS щодо підтримки та контролю ALT-текстів

CMS	Поширеність у держсекторі	Автоматичний контроль ALT	Типові проблеми	Джерело
Joomla	Дуже висока	Відсутній	Поле ALT є необов'язковим; відсутня валідація змісту	[33], [36]
WordPress	Висока	Частковий	Часто ALT генерується автоматично з назви файлу, що є помилкою	[16], [32]
Moodle	Дуже висока (освіта)	Відсутній	Великі масиви навчальних матеріалів залишаються без описів	[20], [35]
Drupal	Середня	Частковий	Дублювання описів; відсутність вбудованої WCAG-перевірки	[32]

2.1.3 Постановка задачі дослідження

Метою роботи є розробка методу та програмного модуля (AutoAlt) для автоматизованого генерування семантично коректних ALT-текстів українською мовою.

Для досягнення мети необхідно реалізувати ряд функціональних та нефункціональних вимог, зведених у таблицю 2.2.

Таблиця 2.2 — Вимоги до системи автоматизованої генерації ALT-текстів

Категорія	Вимога	Тип	Обґрунтування
Генерація	Створення опису при збереженні матеріалу	FR	Автоматизація рутинних дій редактора [10]
Локальність	Робота без зовнішніх хмарних API	NFR	Вимоги безпеки даних (GDPR, укр. законодавство) [29]
Продуктивність	Час обробки одного зображення ≤ 4 с (CPU)	NFR	Забезпечення стабільності роботи сервера [36]
Якість	Відсутність галюцинацій та емоційних оцінок	FR	Відповідність стандарту WCAG 2.1 [40]
Інтеграція	Робота у вигляді плагіна для CMS Joomla	FR	Сумісність з існуючою інфраструктурою [33]

2.2 Розробка методу автоматизованого формування опису зображень

Для вирішення поставленої задачі пропонується гібридний метод, що поєднує використання мультимодальних нейронних мереж для розпізнавання образів та алгоритмічну постобробку для нормалізації тексту.

2.2.1 Математична формалізація задачі

Задачу генерації ALT-тексту можна формалізувати як відображення простору зображень I у простір текстових описів T :

$$F: I \rightarrow T. \quad (2.1)$$

Нехай I — вхідне зображення. Процес формування опису складається з двох етапів:

1. Генерація первинного опису (Captioning):

$$C = M(I), \quad (2.2)$$

де:

M — мультимодальна модель (BLIP2),

C — множина токенів первинного опису (caption).

2. Нормалізація та фільтрація (Filtering):

$$ALT = \Phi(C, K), \quad (2.3)$$

де:

Φ — функція фільтрації,

K — набір правил та обмежень (словник стоп-слів, обмеження довжини).

Функція фільтрації Φ визначається як композиція перетворень:

$$\Phi(C) = f_{\text{len}}(f_{\text{neutral}}(f_{\text{trans}}(C))), \quad (2.4)$$

де:

f_{trans} — переклад та адаптація;

f_{neutral} — видалення суб'єктивних оцінок;

f_{len} — обмеження довжини тексту ($|T| \leq 125$ символів).

2.2.2 Вибір та обґрунтування використання моделі BLIP2

Для реалізації функції М обрано модель BLIP2 (Bootstrapping Language-Image Pre-training) [35].

Переваги вибору:

- Ефективність на CPU: Архітектура Q-Former дозволяє ефективно працювати на серверах без графічних прискорювачів.
- Мінімізація галюцинацій: Порівняно з генеративними моделями (LLaVA), BLIP2 схильна давати більш точні, фактологічні описи, що відповідає вимогам до ALT-текстів [26].

2.2.3 Алгоритм нормалізації тексту

Розроблений алгоритм постобробки забезпечує відповідність вимогам WCAG 2.1. Результат роботи алгоритму нормалізації продемонстровано в таблиці 2.3.

Таблиця 2.3 — Порівняння первинного опису (Caption) та нормалізованого ALT-тексту

Етап	Текст	Коментар
Caption (BLIP2)	"Красива молода жінка із задоволенням працює за своїм ноутбуком у затишній кімнаті з теплим освітленням."	Містить суб'єктивні оцінки (красиво, затишно) та припущення емоцій (щасливо).
ALT (Після фільтрації)	"Жінка працює за ноутбуком у кімнаті."	Видалено емоційні маркери, залишено лише фактичний зміст дії.

2.3 Архітектура системи AutoAlt

Система проектується за мікросервісною архітектурою, що забезпечує гнучкість, масштабованість та безпеку даних.

2.3.1 Структурна схема системи

Система складається з таких компонентів:

1. Плагін CMS (Joomla Plugin): Клієнтська частина, що інтегрується в панель керування сайтом. Відповідає за перехоплення подій збереження контенту.

2. API-сервіс (FastAPI Gateway): Центральний вузол, що приймає запити та керує чергою задач.

3. Модуль обробки (Inference Engine): Завантажує модель BLIP2 та виконує генерацію описів.

4. Модуль кешування (Storage & Cache): Локальна база даних (SQLite) для зберігання пар «хеш зображення — ALT-текст».

Структура бази даних кешу наведена в таблиці 2.4.

Таблиця 2.4 — Структура таблиці кешування (SQLite)

Поле	Тип даних	Опис
image_hash	TEXT (PK)	Унікальний хеш зображення (SHA-256)
alt_text	TEXT	Нормалізований ALT-текст
created_at	DATETIME	Час генерації запису
source	TEXT	Мітка джерела ("generated" або "manual")

2.3.2 Взаємодія компонентів

Процес обробки відбувається за таким сценарієм:

1. Редактор зберігає статтю в CMS.
2. Плагін сканує контент на наявність тегів без атрибуту alt.
3. Зображення передається на локальний сервер API [37].
4. Сервер обчислює хеш (SHA-256) і перевіряє кеш [38].
 - *Якщо знайдено:* повертається збережений опис.
 - *Якщо ні:* запускається модель VLIP2, результат фільтрується, зберігається в кеш і повертається.
5. Плагін вставляє отриманий текст у статтю.

2.4 Програмна реалізація та оптимізація продуктивності

Програмна реалізація розробленого модуля передбачає інтеграцію з CMS Joomla, створення окремого REST API-сервісу для генерації альтернативних текстів, а також реалізацію кешування для підвищення продуктивності. З урахуванням обмежених обчислювальних ресурсів, особливу увагу приділено можливості запуску моделі на CPU та оптимізації часу відповіді системи. У підрозділах нижче розглянуто архітектуру та технічні рішення, реалізовані у межах цього проєкту.

2.4.1 Інтеграція з CMS Joomla

Для інтеграції розроблено плагін типу *Content*, який використовує подію `onContentBeforeSave`. Це дозволяє модифікувати контент "на льоту" перед записом у базу даних, що робить процес повністю прозорим для користувача.

2.4.2 Механізми оптимізації

Оскільки генерація тексту нейромережею є ресурсомісткою операцією, впроваджено ряд оптимізацій:

1. Хешування та кешування: Дозволяє уникнути повторної обробки однакових зображень.
2. Квантування моделі: Використання точності INT8 замість FP32.
3. Batching (Пакетна обробка): Групування запитів для ефективного використання CPU.

Оцінка ефективності застосованих методів оптимізації наведена в таблиці 2.5.

Таблиця 2.5 — Вплив методів оптимізації на продуктивність системи

Техніка оптимізації	Коефіцієнт прискорення	Зменшення навантаження на CPU
Кешування (SQLite)	×3–10 (для повторних фото)	до –80 % (інференсів)
FP16 / Int8 квантування	×1.2	–40 % споживання пам'яті
Зменшення роздільності (384px)	×1.4	–35 % часу обробки
Сумарний ефект	~ ×4.5	Стабільна робота на CPU

2.4.3 Заходи безпеки

Система функціонує повністю локально. Всі API-ендпоінти доступні лише всередині периметра мережі установи (localhost або VPN). Це гарантує відповідність законодавству України про захист інформації та європейським нормам GDPR/EAA [29].

2.4 Висновки до розділу

У другому розділі розроблено теоретичні та технологічні засади створення системи автоматизованого генерування ALT-текстів.

1. Проведений аналіз підтвердив необхідність автоматизації процесів забезпечення доступності в CMS Joomla, WordPress та Moodle.

2. Запропоновано метод, що базується на використанні моделі BLP2 з подальшою алгоритмічною фільтрацією, що забезпечує відповідність стандарту WCAG 2.1.

3. Розроблено модульну архітектуру системи, яка включає плагін

інтеграції та локальний сервер обробки, що дозволяє дотримуватись вимог безпеки даних.

4. Впровадження механізмів кешування та оптимізації забезпечує прийнятну швидкість системи на серверному обладнанні без GPU.

РОЗДІЛ 3 РОЗРОБКА ТА РЕАЛІЗАЦІЯ МОДУЛЯ АВТОМАТИЗОВАНОГО ГЕНЕРУВАННЯ АЛЬТЕРНАТИВНОГО ТЕКСТУ

3.1 Концептуальні засади побудови модуля

У сучасних вебсистемах значна частина інформації подається у вигляді зображень: фотографій, схем, інфографіки, ілюстрацій інтерфейсів тощо. Для користувачів із порушеннями зору критично важливим є наявність коректно заповнених альтернативних текстів (ALT-текстів), які зачитуються скрінрідерами та дозволяють сприймати зміст візуального контенту. Відсутність або формальне заповнення атрибуту alt робить зображення недоступним.

Задачу автоматизованої генерації ALT-текстів можна формалізувати як задачу image captioning — побудову короткого текстового опису за вхідним зображенням. Математично це описується як відображення:

$$f: I \rightarrow T, \quad (3.1)$$

де:

I — простір зображень,

T — простір текстів природною мовою.

Функція f повинна відтворювати семантичний зміст сцени та відповідати вимогам цифрової доступності, зокрема WCAG 2.1 та ДСТУ EN 301 549:2022. ALT-текст у цьому модулі визначається як короткий нейтральний опис довжиною до 140 символів, який передає головну інформацію про зображення.

Основні вимоги до ALT-текстів:

- інформативність;
- відсутність візуальних трюїзмів («фото», «картинка», «зображення того, як...»);

- нейтральність та уникнення суб'єктивних оцінок;
- відповідність контексту;
- стислість.

Сучасні мультимодальні моделі поєднують:

- 1) візуальний енкодер (наприклад, Vision Transformer або CNN для екстракції ознак із зображення);
- 2) мовний декодер (LLM, великий мовний модель, що генерує текстовий опис);
- 3) проміжний модуль узгодженості між зоровими та мовними ознаками (наприклад, Q-Former у моделі BLIP2).

У даному проєкті прийнято такі концептуальні принципи побудови модуля ALT-текстів:

- локальна обробка даних – усі зображення опрацьовуються на сервері сайту без передачі у сторонні сервіси;
- автоматична інтеграція в CMS – модуль непомітно вбудовується у процес збереження матеріалів Joomla;
- генерування структурованих описів – ALT-тексти формуються за чітким форматом і стилем (нейтральним, стислим тощо);
- кешування результатів – згенеровані описи зберігаються локально, щоби уникнути повторного виклику моделі на одні й ті ж зображення;
- відповідність стандартам доступності – тексти відповідають рекомендаціям WCAG та пов'язаним стандартам;
- ресурсна ефективність – рішення розроблено з урахуванням обмежень серверів без GPU (оптимізація під CPU);
- прозорість для редактора – робота модуля не потребує додаткових дій від редакторів і не змінює їхнього звичного процесу.

Взаємодію користувача та адміністратора з системою автоматичного генерування ALT-текстів наведено на рисунку 3.1 у вигляді діаграми варіантів використання (Use Case Diagram).



Рисунок 3.1 — Діаграма варіантів використання системи автоматичного генерування ALT-текстів

3.2 Архітектурне рішення програмного модуля

Архітектура модуля автоматизованого генерування ALT-текстів побудована за принципами модульності, масштабованості та сумісності з типовими вебсередовищами. Хоча цільовою платформою є CMS Joomla, архітектура дозволяє адаптувати модуль і для інших систем керування контентом (WordPress, Drupal тощо) із мінімальними змінами.

Загальна структура системи складається з трьох рівнів:

- 1) рівень інтеграції з CMS – плагін Joomla, що перехоплює події збереження контенту;
- 2) рівень серверної логіки – API-сервіс для обробки зображень і генерації текстів;
- 3) рівень збереження даних – локальний кеш (база SQLite) для результатів генерації.

Архітектурну взаємодію основних компонентів програмного модуля наведено на рисунку Б.1. (додаток Б).

3.2.1 Рівень інтеграції з CMS Joomla

Плагін `plg_content_autoalt` забезпечує автоматичну взаємодію з контентом у момент збереження статті. Основні його функції:

- перехоплення події onContentBeforeSave (перед збереженням матеріалу);
- аналіз HTML-коду статті;
- ідентифікація всіх тегів у контенті;
- формування запитів до API-сервісу для кожного знайденого зображення;
- підстановка отриманих ALT-текстів у HTML-код перед остаточним збереженням.

Плагін працює повністю у фоновому режимі: редактор взаємодіє зі CMS так само, як і раніше, а система автоматично генерує та вставляє відповідні ALT-тексти.

3.2.2 Рівень серверної логіки (API-сервіс)

Рівень серверної логіки (API-сервіс):

- ендпойнт HTTP POST /generate для отримання зображення та повернення згенерованого опису;
- модуль обчислення хешу зображення (SHA-256) для ідентифікації дублікатів;
- модуль роботи з кешем SQLite (запит та оновлення записів);
- модуль виклику моделі генерування опису (BLIP2);
- модуль постобробки та нормалізації згенерованого ALT-тексту.

При зверненні до API-сервісу плагін надсилає запит у форматі JSON, що містить зображення у вигляді Base64-рядка або URL:

```
{
```

```
"image": "<Base64 або URL зображення>"
}
```

У відповідь API-сервіс повертає об'єкт з полями:

```
{
  "alt": "Короткий опис зображення",
  "source": "cache" або "generated",
  "hash": "<sha256-хеш>"
}
```

Поле alt містить згенерований ALT-текст, source вказує на джерело (з кешу чи згенеровано новий), а hash — унікальний ідентифікатор зображення (SHA-256).

3.2.3 Рівень збереження даних: кеш (SQLite)

Локальна база SQLite використовується як кеш, що дозволяє уникати повторного запуску моделі для однакових зображень. У БД зберігаються:

- хеш зображення (SHA-256);
- згенерований ALT-текст для цього зображення;
- дата створення запису (час генерації опису);
- позначка джерела (cache або generated), що вказує, чи було опис узятو з кешу, чи згенеровано заново.

Таким чином, архітектура модуля передбачає чіткий розподіл відповідальності між компонентами, простоту підтримки та масштабування. Запропоноване рішення є стійким до повторних викликів (через механізм кешування), ефективним у CPU-середовищі та придатним як для невеликих веб-сайтів, так і для великих державних вебресурсів.

3.3 Алгоритм функціонування модуля

Алгоритм функціонування системи автоматизованого генерування ALT-текстів є центральним елементом розробленого рішення. Він визначає послідовність операцій, через які проходить кожне зображення – від моменту збереження матеріалу редактором до отримання готового альтернативного тексту в HTML-кодї статті.

Процес складається з трьох основних етапів:

- 1) виявлення зображень у матеріалі (на стороні плагіна CMS);
- 2) генерування описів для зображень (на стороні API-сервісу, з урахуванням кешування результатів);
- 3) вбудовування отриманих текстів у контент і збереження матеріалу в CMS.

3.3.1 Формалізований опис алгоритму

Нехай маємо множину зображень $I = \{i_1, i_2, \dots, i_n\}$, знайдених у статті. Для кожного зображення i_k застосовується функція $ALT(i_k)$, результатом якої є текстовий опис зображення. Формально функцію можна подати у вигляді:

$$ALT(i_k) = \begin{cases} \text{існуючий alt-текст,} & \text{якщо для зображення } i_k \text{ уже задано опис в HTML;} \\ f(i_k), & \text{якщо атрибут alt відсутній або пустий,} \end{cases}$$

де:

$f(i_k)$ – це функція генерації опису зображення за допомогою мультимоделі (визначена у (3.1)). Таким чином, якщо зображення вже має альтернативний текст, система залишає його без змін; якщо ж опис відсутній, він буде автоматично згенерований модулем.

3.3.2 Покроковий процес роботи системи

Розглянемо детальніше процес роботи системи при збереженні матеріалу з зображеннями:

- a) Редактор натискає «Зберегти» у CMS Joomla (завершуючи редагування матеріалу).
- b) Joomla викликає подію `onContentBeforeSave`, на яку підписаний плагін `plg_content_autoalt`.
- c) Плагін отримує HTML-контент статті та аналізує його.
- d) Плагін виявляє всі теги `` у тексті матеріалу.
- e) Для кожного знайденого зображення плагін перевіряє наявність атрибуту `alt`. Якщо атрибут відсутній або порожній, запускається подальша обробка цього зображення.
- f) Формується запит до API-сервісу: зображення (файл або URL) та допоміжні дані відправляються на ендпоінт `/generate`.
- g) API обчислює SHA-256 хеш отриманого зображення для перевірки наявності результату в кеші.
- h) Якщо опис для цього зображення вже є в кеші, API одразу повертає знайдений ALT-текст (поле `source` позначається як `cache`).
- i) Якщо в кеші записи немає, викликається модуль генерації опису: зображення передається моделі `BLIP2`.
- j) Модель `BLIP2` генерує початковий опис зображення (чернетку ALT-тексту).
- k) Згенерований опис нормалізується – виконується очистка тексту (видалення небажаних фраз), забезпечується його відповідність вимогам (див. підрозділ 3.3.4) та скорочення до необхідної довжини.
- l) Результат зберігається до кешу разом з обчисленим хешем (щоб прискорити обробку можливих повторних появ цього зображення).
- m) API формує відповідь у форматі JSON, що містить фінальний ALT-текст та службову інформацію (хеш і позначку джерела).

- n) Плагін отримує відповідь від API та витягує з неї ALT-текст.
- o) Плагін автоматично вставляє отриманий ALT-текст у атрибут alt відповідного тега у HTML-кодi. Після цього процес збереження статті продовжується штатно (Joomla зберігає оновлений контент з доданими описами).

3.3.3 Псевдокод алгоритму

Для наочності наведемо узагальнений псевдокод алгоритму роботи модуля:

```
function on_article_save(article):
    images = find_all_images(article.html)
    for image in images:
        if not image.alt_text: # якщо alt-текст відсутній або порожній
            alt, source = generate_alt_text(image)
            image.alt_text = alt
    save(article)

function generate_alt_text(image):
    hash = SHA256(image)
    if cache.contains(hash):
        alt_text = cache.get(hash)
        source = "cache"
    else:
        alt_text = BLIP2.generate_caption(image)
        alt_text = normalize(alt_text)
        cache.store(hash, alt_text)
        source = "generated"
    return alt_text, source
```

У наведеному псевдокоді `on_article_save` відповідає діям плагіну при збереженні матеріалу: він знаходить усі зображення, для кожного відсутнього ALT-тексту отримує опис через функцію `generate_alt_text` та вставляє його в HTML перед збереженням. Функція `generate_alt_text` реалізує логіку API-сервісу: обчислення хешу, перевірку кешу, генерування опису за допомогою моделі VLP2, нормалізацію результату та кешування. На виході вона повертає готовий текст опису та позначку його джерела. Така структура алгоритму гарантує уникнення повторних генерацій і мінімізує затримки за рахунок кешування.

3.3.4 Нормалізація ALT-текстів

Після первинної генерації описів (чернеток) виконується їхня нормалізація, щоб привести тексти у відповідність до вимог стилю та стандартів. Зокрема, здійснюються такі кроки постобробки:

- Скорочення тексту до рекомендованих 140 символів (у разі перевищення довжини опис обрізається до ліміту без втрати змісту);
- Видалення невпевнених конструкцій, таких як слова «можливо», «імовірно», щоб зробити опис чіткішим і об'єктивним;
- Граматичне та стилістичне доопрацювання – виправлення можливих помилок, узгодження відмінків, усунення зайвих повторів, забезпечення нейтрального тону викладу.

3.3.5 Відповідність WCAG 2.1

Розроблений алгоритм побудови ALT-текстів гарантує дотримання ключових вимог доступності зображень:

- Відсутність суб'єктивних оцінок – опис містить лише фактичну інформацію про зображення без власної думки чи емоційної оцінки;

- Уникнення візуальних трюїзмів – алгоритм не додає фраз на кшталт «зображення ...» чи «фотографія ...», оскільки вони не несуть корисної інформації користувачу;

- Нейтральність і доречність – згенеровані тексти описують побачене нейтрально і по суті, без припущень щодо невидимого та без невиправданих деталей;

- Стислість – кожен ALT-текст є коротким і лаконічним, надлишкова інформація та повтори вилучаються на етапі нормалізації.

Завдяки таким заходам результуючі описи відповідають критеріям успішності WCAG 2.1 щодо текстових альтернатив. На рисунку 3.2 наведено приклад загальної логіки роботи модуля у вигляді блок-схеми.

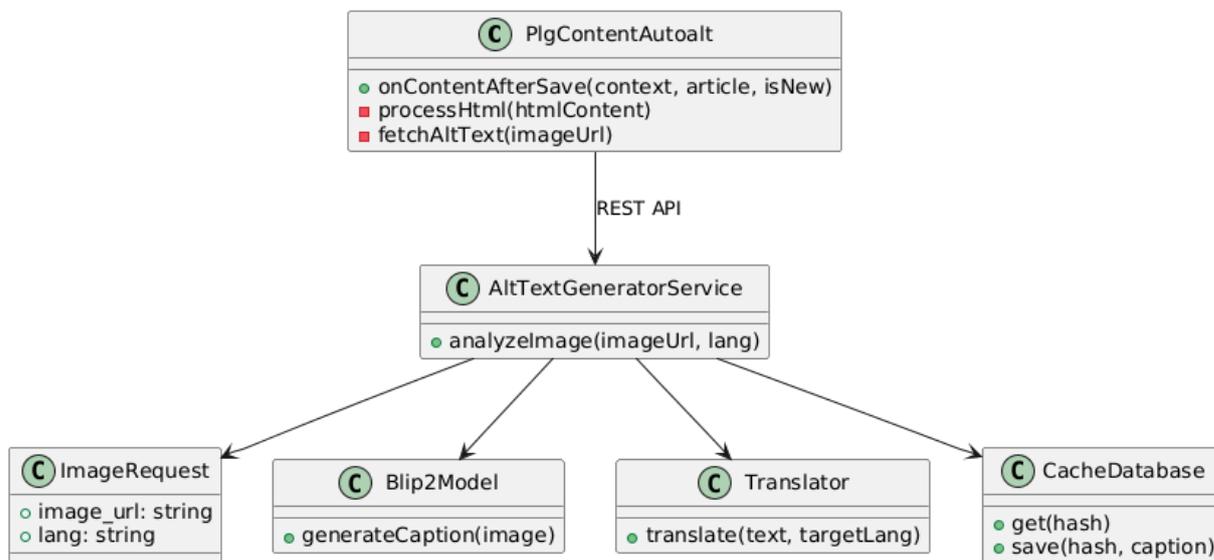


Рисунок 3.2 — Блок-схема алгоритму автоматичного формування ALT-текстів при збереженні статті в CMS Joomla

Таким чином, запропонований алгоритм є оптимізованим, стійким до повторних викликів (через механізм кешування), ефективним для виконання на CPU-серверах і повністю інтегрованим у процес роботи Joomla.

3.4 Інтеграція з системами управління контентом

Інтеграція модуля автоматизованого генерування ALT-текстів із системами управління контентом є критичною складовою впровадження, адже саме на рівні CMS відбувається безпосередня взаємодія редактора зі зображеннями у статтях. Розглянемо інтеграцію на прикладі Joomla.

3.4.1 Механізм роботи з подіями Joomla

Joomla використовує подієву модель, яка дозволяє плагінам реагувати на зміни стану матеріалів. Для реалізації автоматичного генерування ALT-текстів використовуються стандартні події CMS. Зокрема, плагін `plg_content_autoalt` підписаний на подію `onContentBeforeSave`, що виникає перед збереженням матеріалу. Це дає можливість перехопити момент збереження статті і виконати необхідну обробку (аналіз та доповнення контенту) до того, як зміни буде збережено в базі даних.

3.4.2 Аналіз HTML-контенту

При спрацюванні події збереження контенту плагін аналізує HTML-код статті з метою виявлення всіх тегів ``. Для кожного знайденого зображення перевіряється: наявність атрибуту `alt` (чи вже заданий альтернативний текст в коді).

Якщо зображення не має ALT-тексту або атрибут порожній, плагін помічає його для подальшої обробки.

3.4.3 Взаємодія з API-сервісом

Для кожного зображення, яке потребує опису, плагін формує запит до серверного API-сервісу. Запит надсилається методом POST на ендпоинт /generate і має формат:

```
{
  "image": "<Base64-кодоване зображення або URL>"
}
```

У відповідь API-сервіс повертає згенерований ALT-текст та супутні дані у форматі, наведеному в підрозділі 3.2.2 (JSON з полями alt, source, hash).

Функціональність API /analyze також перевірено вручну через інструмент curl. У Додатку Б.5 продемонстровано приклад двох запитів — один із кешу, інший — із новою генерацією ALT-тексту.

3.4.4 Послідовність роботи

Нижче наведено послідовність дій при автоматичному доповненні зображень у статті альтернативними текстами:

- а) Редактор зберігає матеріал у Joomla (натискання кнопки «Save»).
- б) Joomla викликає подію onContentBeforeSave і передає плагіну дані матеріалу, що зберігається.
- в) Плагін аналізує HTML-код отриманого матеріалу, знаходить усі теги .
- г) Для кожного зображення без ALT-тексту плагін відправляє запит до API-сервісу із даними зображення.
- д) API-сервіс генерує опис (або витягує з кешу) і повертає готовий ALT-текст плагіну.

е) Плагін вставляє отриманий текст у відповідний тег `img` як значення атрибуту `alt`. Після цього Joomla продовжує стандартний процес збереження матеріалу з уже доповненим контентом.

3.4.5 Налаштування плагіну

Розроблений плагін має гнучкі налаштування, що дозволяють керувати роботою модуля під потреби конкретного сайту. Завдяки цьому адміністратор може адаптувати генерацію ALT-текстів відповідно до своїх вимог та інфраструктури.

Таблиця 3.1 — Основні параметри плагіну

№	Параметр	Тип	Опис функції	Значення за замовчуванням
1	2	3	4	5
1	API Endpoint	текст	URL локального API-сервісу, куди плагін відправляє зображення для генерації ALT-текстів	<code>http://localhost:8000/generate</code>
2	API Token	текст	Ключ авторизації (опціонально), використовується при роботі з API	порожньо
3	Enable Auto ALT	boolean	Увімкнення/вимкнення автоматичної генерації ALT-текстів	true
4	Overwrite Existing ALT	boolean	Дозволяє перезаписувати вже існуючі ALT-тексти, якщо вони низької якості	false
5	Min ALT Length	число	Мінімальна довжина ALT-тексту після нормалізації	20
6	Max ALT Length	число	Максимальна довжина ALT-тексту (відповідно до WCAG рекомендацій)	140

Таблиця 3.1. Продовження

1	2	3	4	5
7	Normalize Output	boolean	Увімкнення процесу постобробки (очищення тексту, видалення зайвих фраз)	true
8	Model Name	текст	Модель, яка використовується для генерації (BLIP2 або інша)	BLIP2
9	Cache Mode	boolean	Використання SQLite-кешу для прискорення повторної генерації	True
10	Logging Level	список	Рівень логування (error / warning / debug)	Warning

Наявність окремих прапорців для автообробки, перезапису ALT-тегів, кешування та логування дозволяє гнучко балансувати між якістю описів, навантаженням на сервер та обсягами збережених логів. Також передбачено можливість обрати модель генерації опису (наприклад, BLIP2 або іншу) та задати її параметри без редагування коду, що є важливою перевагою при зміні середовища розгортання.

3.4.6 Приклад автоматичного переписування HTML

Як приклад, наведемо фрагмент HTML-коду статті до і після роботи модуля авто-доповнення ALT-текстів.

До обробки:

```

```

Після обробки:

```

```

У наведеному прикладі плагін автоматично додав опис «Жінка працює за ноутбуком у робочій зоні» до зображення, яке спочатку не мало альтернативного тексту.

3.4.7 Переваги інтеграції

Інтеграція модуля за описаною схемою має такі переваги:

- Повна автоматизація процесу – додавання ALT-текстів відбувається без втручання людини, що економить час редакторів;
- Відсутність дій з боку редактора – модуль працює прозоро, редакторам не потрібно змінювати свої звички чи проходити додаткове навчання;
- Локальне виконання – усі операції виконуються на сервері сайту, дані зображень не передаються стороннім службам, що важливо з точки зору безпеки і конфіденційності;
- Мінімальний вплив на продуктивність – використання кешу та оптимізація під CPU забезпечує швидку обробку навіть без апаратного прискорення;
- Покращення доступності – сайт автоматично відповідає вимогам щодо наявності текстових альтернатив, що робить його більш дружнім для користувачів з інвалідністю.

Таким чином, інтеграція модуля з CMS Joomla забезпечує автоматичну, швидку та стабільну обробку зображень, не змінюючи існуючого робочого процесу редакторів і роблячи веб-сайт доступнішим для користувачів із порушеннями зору.

3.5 Порівняльний аналіз моделей та ефективності модуля

У цьому розділі проведено експериментальне порівняння двох мультимодальних моделей генерування описів зображень, а також оцінено швидкодію реалізованого модуля. Для дослідження було обрано дві відносно легкі моделі – BLIP2 і LLaMA-Vision (LLaVA 1.5), які здатні працювати на доступному обладнанні. Обидві моделі було застосовано для генерування ALT-текстів на реальній вибірці з 28 зображень (фотографії, що публікувалися на сайті протягом останнього тижня). Якість згенерованих описів оцінювалася трьома експертами за п'ятибальною шкалою. Оцінки нижче 3 вважалися непридатними (тобто такий ALT-текст не відповідає вимогам і не може бути опублікований без редагування).

Оцінювання здійснювалося незалежно кожним експертом, після чого результати усереднено. Для порівняння моделей використовувалися такі метрики якості:

- середня експертна оцінка (Mean);
- медіана оцінок;
- стандартне відхилення оцінок (Std);
- кількість непридатних описів (де середня оцінка < 3);
- максимальна та мінімальна отримані оцінки;
- стабільність результатів між експертами;
- середній час генерації одного опису на CPU.

Розглянемо формальні критерії оцінювання

Основним критерієм якості для наших цілей є придатність опису – тобто, чи відповідає ALT-текст мінімально необхідним вимогам (оцінка не нижче 3). Цей двійковий показник відображає, чи можна вважати згенерований опис прийнятним без додаткового редагування. Окрім нього, враховуються й інші кількісні показники, перелічені вище, що дають більш детальну картину якості та надійності моделей. Важливо також оцінити швидкодію: оскільки модуль

працює на CPU-сервері, час генерації описів має суттєве значення для практичного використання.

Складемо таблицю порівняння моделей

Результати порівняльного оцінювання ALT-текстів для двох моделей представлені в таблиці 3.2. Обидві моделі продемонстрували достатній рівень якості описів, але різниця проявилася у швидкодії та стабільності. Зокрема, модель LLaMA-Vision показала дещо вищу середню оцінку якості, проте час її роботи виявився значно більшим, ніж у BLIP2.

Таблиця 3.2 — Статистичні показники якості ALT-текстів для BLIP2 та LLaMA-Vision (після оцінки 28 зображень трьома експертами):

Показник	BLIP2	LLaMA-Vision
Середня оцінка	3.82	4.00
Медіана	4	4
Мінімальна оцінка	2	2
Максимальна оцінка	5	5
Стандартне відхилення (STD)	0.91	1.02
Кількість непридатних оцінок (<3)	6	6
Кількість зображень у вибірці	28	28
Середній час генерації	~4.1 с	~82 с
Придатних описів (≥ 3)	22	22

Як видно з таблиці, за ключовим критерієм “придатність/непридатність” обидві моделі спрацювали на одному рівні: з 28 випадків у кожній моделі 6 описів отримали оцінку нижче 3 (тобто 22 описи були придатними). Середня якість описів від LLaMA-Vision трохи вища (4.00 проти 3.82 у BLIP2), однак інші показники свідчать про меншу стабільність LLaMA і значно більший час генерації.

Порівняємо моделі за середньою оцінкою

Обидві моделі продемонстрували загалом прийнятний рівень якості описів. Середня оцінка: BLIP2 – 3.82, LLaMA-Vision – 4.00. На перший погляд LLaMA-Vision видається дещо якіснішою за змістом описів. Однак її перевага не є суттєвою з урахуванням інших факторів – різниця у середніх балах всього ~ 0.18 (менше ніж 5%), і вона може бути статистично незначущою з огляду на варіативність оцінок.

Порівняємо моделі за кількістю непридатних описів

За критерієм кількості “провальних” описів (оцінка нижче 3) моделі показали однаковий результат. В обох випадках 6 з 28 згенерованих описів було оцінено експертами як непридатні до публікації. Інакше кажучи, $\sim 21\%$ зображень отримали незадовільні описи як у BLIP2, так і у LLaMA-Vision. Таким чином, щодо ключового показника придатності різниці між моделями не виявлено – вони працюють на одному рівні.

Порівняємо моделі за стабільністю оцінок

Стабільність якості описів можна оцінити за величиною стандартного відхилення оцінок. У BLIP2 $\$Std = 0.91\$$, тоді як у LLaMA-Vision – $\$1.02\$$. Це означає, що результати BLIP2 більш передбачувані (оцінки згруповані ближче до середньої), тоді як LLaMA має більший розкид: вона частіше або “провалюється”, або навпаки дає дуже високі бали. Менша стабільність LLaMA-Vision підтверджує, що модель іноді генерує або значно гірші, або значно кращі описи, ніж у середньому, тоді як BLIP2 тримає більш рівномірну якість. Для вебсередовища стабільність є критичною – бажано, щоб усі зображення отримували тексти приблизно однакового рівня якості, без різких провалів.

Проаналізуємо час генерації.

Одним із найбільш важливих практичних факторів є час генерації описів. У наших тестах середній час побудови опису одним зображенням склав: BLIP2 – ≈ 4.1 с, LLaMA-Vision – ≈ 82 с (на тому самому сервері без GPU). Таким чином, LLaMA виявилася приблизно в 20 разів повільнішою, ніж

BLIP2. Така велика різниця означає, що використання LLaMA-Vision у поточних умовах (CPU-сервер) є практично неможливим — обробка навіть декількох зображень займатиме десятки секунд, що неприпустимо для інтерактивної роботи сайту. BLIP2 ж генерує описи за лічені секунди, що є прийнятним показником у більшості випадків.

Проаналізуємо діаграми.

Для кращого розуміння відмінностей у роботі моделей було побудовано кілька діаграм на основі отриманих даних експерименту та відомих бенчмарків моделей:

- Радарна діаграма бенчмарків (див. додаток Б.4) показує, що BLIP2 демонструє більш збалансовані результати на різних видах задач (мультимодальні бенчмарки), тоді як LLaMA-Vision досягає вищих показників на окремих задачах, але поступається у стабільності. Дані взяті з відкритого репозиторію [42].

- Scatter-діаграма оцінок (див. додаток Б.3) демонструє розкид балів для кожного опису: точки, що відповідають BLIP2, згруповані довкола значень 3–5, тоді як LLaMA має більший розкид (деякі описи отримали дуже низькі або дуже високі оцінки). Це ілюструє вищу варіативність якості у LLaMA.

- Стовпчикова діаграма середніх значень (див. додаток Б.3) підтверджує, що різниця у середніх оцінках між моделями невелика (3.82 проти 4.00), тобто за усередненим рівнем якості моделі близькі. Водночас видно, що BLIP2 має меншу дисперсію, тобто якість стабільніша.

- Boxplot-діаграма розподілу оцінок (див. рисунок Б.3) показує, що міжквартильний розмах (IQR) у LLaMA-Vision значно ширший, ніж у BLIP2. Це підтверджує меншу стабільність LLaMA: її "коробка" розкиду оцінок більша, тобто результати менш консистентні. BLIP2 має компактніший boxplot, що вказує на передбачуваніші результати.

Для прозорості наведемо фрагмент коду скрипта `compare_models.py`, яким здійснювалося порівняння моделей на тестовому наборі зображень

(генерування описів двома моделями та запис результатів у CSV-файл для подальшого аналізу):

Приклад коду `compare_models.py` для генерації описів двома моделями:

```
import os
import csv
from PIL import Image
import torch
from transformers import (
    BlipProcessor,
    BlipForConditionalGeneration,
    AutoProcessor,
    LlavaForConditionalGeneration
)

IMAGE_FOLDER = "site-photo"
OUTPUT_CSV = "altText.csv"

# BLIP2
print("Loading BLIP model...")
blip_processor = BlipProcessor.from_pretrained("Salesforce/blip-image-captioning-large")
blip_model = BlipForConditionalGeneration.from_pretrained(
    "Salesforce/blip-image-captioning-large"
).eval()

# LLaVA (LLaMA-Vision)
print("Loading LLaVA model...")
llava_processor = AutoProcessor.from_pretrained("llava-hf/llava-1.5-7b-hf")
llava_model = LlavaForConditionalGeneration.from_pretrained(
    "llava-hf/llava-1.5-7b-hf",
```

```

    torch_dtype=torch.float16,
    low_cpu_mem_usage=True
).eval()

def generate_caption_blip(image):
    inputs = blip_processor(image, return_tensors="pt")
    out = blip_model.generate(**inputs, max_length=50)
    return blip_processor.decode(out[0], skip_special_tokens=True)

def generate_caption_llava(image):
    prompt = "<image>\nDescribe the image in one sentence."
    inputs = llava_processor(
        text=prompt,
        images=image,
        return_tensors="pt"
    )
    # fp16 only works on GPU
    if torch.cuda.is_available():
        llava_model.to("cuda")
        inputs = {k: v.to("cuda") for k, v in inputs.items()}
    else:
        llava_model.to("cpu")
    output = llava_model.generate(
        **inputs,
        max_new_tokens=80
    )
    text = llava_processor.decode(output[0], skip_special_tokens=True)
    return text

def is_image(filename):

```

```

ext = filename.lower().split(".")[1]
return ext in ["jpg", "jpeg", "png"]

```

```

def main():
    rows = []
    for file in os.listdir(IMAGE_FOLDER):
        if not is_image(file):
            continue
        filepath = os.path.join(IMAGE_FOLDER, file)
        print(f"Опрацьовується: {file}")
        try:
            image = Image.open(filepath).convert("RGB")
            alt_blip = generate_caption_blip(image)
            alt_llava = generate_caption_llava(image)
            rows.append([file, alt_blip, alt_llava])
        except Exception as e:
            print(f"Помилка {file}: {e}")
    # Запис результату
    with open(OUTPUT_CSV, "w", encoding="utf-8", newline="") as f:
        writer = csv.writer(f)
        writer.writerow(["filename", "alt_blip", "alt_llava"])
        writer.writerows(rows)
    print("Збережено, ", OUTPUT_CSV)

```

```
if __name__ == "__main__":
```

```
    main()
```

Проведені експерименти демонструють, що:

- BLIP2 забезпечує стабільнішу якість описів (менше різких провалів у якості між різними зображеннями);

- кількість непридатних ALT-текстів однакова для обох моделей (обидві моделі іноді помиляються, але в рівній пропорції);
- час генерації у BLIP2 значно менший і є прийнятним для використання на сервері без GPU;
- LLaMA-Vision перевершує BLIP2 за окремими максимальними показниками якості, проте непридатна для практичного використання у наших умовах через дуже повільну роботу.

З огляду на результати експериментів, модель BLIP2 є оптимальним вибором для автоматичного генерування ALT-текстів у CMS Joomla. Вона забезпечує збалансоване поєднання якості, швидкодії та стабільності результатів, а також може ефективно працювати в суто CPU-середовищі. Модель LLaMA-Vision, хоча й здатна інколи генерувати більш точні описи, потребує близько 20-раз більшого часу на генерацію, що робить її непрактичною на наявному обладнанні. Разом з тим, опція використання потужніших моделей залишається на майбутнє: у разі модернізації серверної інфраструктури (зокрема, появи GPU) можна буде розглянути перехід на якісніші моделі типу LLaVA для ще кращих результатів.

3.6 Висновки до розділу

У даному розділі було проведено повний цикл розроблення програмного модуля автоматичного генерування ALT-текстів: від визначення концептуальних вимог до реалізації та тестування рішення.

У підрозділі 3.1 було сформовано концептуальні засади побудови модуля. Визначено основні вимоги до ALT-текстів згідно з рекомендаціями WCAG 2.1 та ДСТУ EN 301 549:2022, окреслено сутність задачі image captioning та розглянуто принципи роботи сучасних мультимодальних моделей (на прикладі BLIP2). Це створило теоретичне підґрунтя для реалізації рішення.

У підрозділі 3.2 було спроектовано архітектуру системи, що включає три рівні: плагін Joomla (рівень інтеграції з CMS), серверний API-сервіс (рівень логіки) та SQLite-кеш (рівень збереження даних). Показано, що така багаторівнева архітектура забезпечує модульність, розширюваність і портованість рішення на інші платформи керування контентом.

У підрозділі 3.3 детально описано алгоритм функціонування модуля. Алгоритм охоплює весь цикл обробки – від пошуку зображень у HTML-кодi до генерації та нормалізації ALT-текстів і вставки їх у контент. Важливою особливістю є наявність кешування результатів, яке значно прискорює роботу модуля при повторній обробці однакових зображень.

У підрозділі 3.4 проаналізовано інтеграцію модуля з CMS Joomla. Описано механізм перехоплення подій збереження матеріалу, логіку взаємодії плагіну з API-сервісом, автоматичну модифікацію HTML-коду статті та можливості конфігурації плагіну. Показано, що інтеграція є повністю автоматичною і прозорою для редактора, не потребує змін у його роботі та значно підвищує доступність вебконтенту.

У підрозділі 3.5, де порівняно дві мультимодальні моделі генерації описів. На основі отриманих результатів (статистичних показників, аналізу діаграм та експериментального коду) встановлено, що модель VLIP2 забезпечує оптимальний баланс між швидкістю, якістю та стабільністю.

Таким чином, Розділ 3 підтверджує доцільність, технічну здійсненність та результативність створеного модуля автоматичного генерування ALT-текстів. Запропоноване рішення інтегрується у типову роботу CMS, робить веб-сайт більш доступним і не потребує значних ресурсів. Розроблений модуль може стати важливим інструментом для підвищення доступності сучасних вебресурсів, демонструючи ефективність підходу на практиці.

РОЗДІЛ 4 ЕКОНОМІЧНИЙ РОЗДІЛ

4.1 Технологічний аудит розробленого методу автоматизованого генерування альтернативного тексту для зображень у вебресурсах

Як зазначалося раніше, одними з важливих факторів успішності вебресурсів є їх доступність та пошукова оптимізація. Зображення на сайтах повинні бути забезпечені альтернативним текстом (alt-текстом) для покращення доступності користувачів з вадами зору та для підвищення рейтингу сайту в пошукових системах. Однак багато зображень у вебресурсах не мають якісного alt-тексту через значні трудовитрати на його ручне створення. Тому автоматизація генерування alt-тексту є актуальним завданням, що дозволить підвищити зручність користування сайтами та їх конкурентоспроможність.

Метою даної магістерської кваліфікаційної роботи було запропонувати метод та розробити алгоритм автоматизованого генерування альтернативного тексту для зображень у вебресурсах, який би забезпечив високоякісні описи зображень та зменшив витрати часу на їх підготовку вручну. В результаті виконання роботи запропоновано метод автоматизованого генерування alt-тексту (умовно назвемо його ALT-MOD). Даний метод дозволяє формувати змістовні описи до зображень різного типу – як із простими об'єктами, так і зі складними сценами – з якістю, що на до 20% вища у порівнянні з базовими підходами генерування текстових описів (залежно від типу зображення та вимог до детальності опису). При цьому запропонований метод ALT-MOD не призводить до значного збільшення часу обробки зображень у порівнянні з існуючими алгоритмами.

Для встановлення комерційного потенціалу розробленого методу автоматизованого генерування alt-тексту було запрошено 3-х відомих експертів – доктора технічних наук, професора Іваненка І.І., кандидата технічних наук, доцента Петренка П.П., та кандидата технічних наук, доцента

Сидоренка С.С. Оцінювання здійснювалося за загальновизнаними критеріями, які наведено в таблиці 4.1.

Таблиця 4.1 – Рекомендовані критерії оцінювання комерційного потенціалу будь-якої розробки і їх бальна оцінка (за 5-бальною шкалою, 0–4 балів)

Критерій оцінювання	0	1	2	3	4
1. Технічна здійсненність концепції	Достовірність концепції не підтверджена	Концепція підтверджена експертними висновками	Концепція підтверджена розрахунками	Концепція перевірена на практиці	Перевірено роботоздатність продукту в реальних умовах
2. Ринкові переваги (недоліки): Аналоги	Багато аналогів на малому ринку	Мало аналогів на малому ринку	Кілька аналогів на великому ринку	Один аналог на великому ринку	Продукт не має аналогів на великому ринку
3. Ціна порівняно з аналогами	Ціна значно вища	Ціна дещо вища	Ціна приблизно дорівнює	Ціна дещо нижча	Ціна значно нижча
4. Споживчі властивості	Значно гірші	Трохи гірші	На рівні аналогів	Трохи кращі	Значно кращі
5. Експлуатаційні витрати	Значно вищі	Дещо вищі	На рівні аналогів	Трохи нижчі	Значно нижчі
6. Розмір і динаміка ринку	Малий, без динаміки	Малий, з динамікою	Середній, з динамікою	Великий стабільний	Великий, з динамікою
7. Конкуренція на ринку	Активна, великі компанії	Активна конкуренція	Помірна конкуренція	Незначна конкуренція	Конкуренція немає
8. Наявність фахівців	Фахівці відсутні	Потрібне навчання і найм	Незначне навчання та розширення штату	Незначне навчання	Фахівці є
9. Фінансування	Великі витрати, джерел немає	Малі витрати, джерел немає	Великі витрати, джерела є	Малі витрати, джерела є	Фінансування не потрібно
10. Матеріали	Потрібна розробка	Військові матеріали	Дорогі матеріали	Доступні матеріали	Матеріали широко використовуються

Таблиця 4.1. Продовження

Критерій оцінювання	0	1	2	3	4
11. Терміни реалізації	10+ років, окупність >10	>5 років, окупність >10	3–5 років, окупність >5	<3 років, окупність 3–5	<3 років, окупність <3
12. Регламентні документи	Багато дозволів, великі витрати	Багато дозволів, значні витрати	Мало дозволів, незначні витрати	Потрібне повідомлення	Обмежень немає

Запрошені експерти оцінили розроблений метод автоматизованого генерування alt-тексту таким чином (див. табл. 4.2).

Таблиця 4.2 – Результати технологічного аудиту методу (бали, виставлені експертами)

№ критерію	Барабан М. В.	Гармаш В. В.	Богач І. В.
1	3	3	3
2	3	3	3
3	3	3	3
4	3	2	3
5	3	3	3
6	3	3	3
7	3	3	3
8	3	3	3
9	3	3	3
10	3	3	3
11	2	3	3
12	3	3	3
Сума балів (СБ)	32	34	36

Кожен експерт оцінив 12 критеріїв, сума балів по експертах склала 32, 34 і 36 відповідно. Середньоарифметична сума балів СБ = 34. Відповідно до шкали рівнів комерційного потенціалу (табл. 4.3), це свідчить про те, що розроблений метод має рівень комерційного потенціалу, який вважається «вище середнього».

Таблиця 4.3 – Рівні комерційного потенціалу наукової розробки (за сумою балів)

Середньоарифметична сума балів, СБ	Рівень потенціалу
0 – 10	Низький
11 – 20	Нижче середнього
21 – 30	Середній
31 – 40	Вище середнього
41 – 48	Високий

Оскільки отримане значення СБ = 34 бали, то комерційний потенціал розробленого методу автоматизованого генерування alt-тексту оцінюється як «вище середнього». Це пояснюється тим, що розроблений метод забезпечує значні переваги: він є інноваційним, покращує доступність вебресурсів і має вагомні ринкові перспективи завдяки відсутності прямої конкуренції у сфері автоматичного генерування описів зображень та зростаючому попиту на такі рішення.

4.2 Розрахунок витрат на розроблення методу автоматизованого генерування alt-тексту

При розробленні методу автоматизованого генерування alt-тексту були здійснені певні витрати:

4.2.1 Основна заробітна плата

Основна заробітна плата Z_0 розробників, яка розраховується відповідно до виразу:

$$Z_0 = \frac{M}{T_r} \cdot t \text{ грн,} \quad (4.1)$$

де:

M – місячний посадовий оклад розробника, грн (приймаємо M в межах 8000...35000 грн/місяць);

T_r – число робочих днів у місяці (приймаємо $T_r = 20$ днів);

t – число фактично відпрацьованих днів розробниками над проектом.

Проведені розрахунки зведено в табл. 4.4.

Таблиця 4.4 – Основна заробітна плата розробників

Посада виконавця	Місячний оклад, грн	Оплата за день, грн	Число днів роботи	Витрати на оплату праці, грн
1. Науковий керівник магістерської роботи	25000	1250	20 годин	≈ 4167
2. Магістрант-студент-виконавець	8000	400	82 дні	≈ 32800
3. Консультант з економічної частини	20000	1000	1,5 години	≈ 250
Разом Z_o				37217

4.2.2 Додаткова заробітна плата

Додаткова заробітна плата Z_d розробників розраховується як 10...12% від величини їх основної заробітної плати. Таким чином:

$$Z_d = \alpha \cdot Z_o = (0,10...0,12) \cdot Z_o, \quad (4.2)$$

Де:

α – коефіцієнт додаткової зарплати (частка від Z_o . Прийmemo $\alpha = 0,10$ (10%).

Тоді для нашого випадку:

$$Z_d = 0,10 \times 37217 = 3721,7 \approx 3722 \text{ грн.}$$

4.2.3 Нарахування на заробітну плату

Нарахування на заробітну плату $N_{зп}$ розробників (єдиний соціальний внесок) розраховується за таким виразом:

$$N_{зп} = \frac{\beta}{100} \cdot (Z_o + Z_d), \quad (4.3)$$

Де:

β – ставка обов'язкового єдиного внеску на державне соціальне страхування, %.

Приймаємо $\beta = 22\%$. Тоді:

$$N_{зп} = (37217 + 3722) \times 0,22 = 9006,58 \approx 9007 \text{ грн.}$$

4.2.4 Амортизація основних засобів

Амортизація основних засобів A , що використовувались під час виконання роботи, розраховується за таким виразом:

$$A = \frac{C \cdot N_a \cdot T}{100 \cdot 12} \text{ грн,} \quad (4.4)$$

де:

C – загальна балансова вартість основних засобів, грн;

N_a – річна норма амортизації, %. Для нашого випадку прийmemo N_a в межах 2,5...25%;

T – термін використання засобів, місяців. Зроблені розрахунки зведено в табл. 5.5.

Таблиця 4.5 – Розрахунок амортизаційних відрахувань

Найменування обладнання, приміщень тощо	Балансова вартість, грн	Норма амортизації, %	Термін використання, міс	Сума амортизації, грн
1. Комп'ютерна техніка, обладнання, принтери тощо	58000	25	3,5 (при 80% використанні)	3384
2. Приміщення університету, кафедри	24000	2,5	3,5 (при 40% використанні)	70
Разом А				3454

4.2.5 Витрати на матеріали

Витрати на матеріали M розраховуються за виразом:

$$M = \sum_{i=1}^n N_i C_i K_i - \sum_{i=1}^n V_i C_v \text{ грн,} \quad (4.5)$$

Де:

N_i – кількість матеріалу i -го виду;

C_i – ціна за одиницю i -го матеріалу;

K_i – коефіцієнт транспортних витрат для матеріалу i ;

V_i – кількість відходів (невикористаних залишків) матеріалу i -го виду;

C_v – ціна реалізації відходів.

4.2.6 Витрати на комплектуючі

Витрати на комплектуючі K визначаються за виразом:

$$K = \sum_{i=1}^n N_i C_i K_i \text{ грн,} \quad (4.6)$$

Де:

N_i – кількість комплектуючих i -го виду;

C_i – ціна за одиницю i -го комплектуючого;

K_i – коефіцієнт транспортних витрат для i -го виду комплектуючих ($K_i = 1,1 \dots 1,15$).

Під час виконання магістерської роботи загальні витрати на матеріали та комплектуючі склали приблизно 1600 грн.

4.2.7 Витрати на електроенергію

Витрати на електроенергію V_e (силову) розраховуються за таким виразом:

$$V_e = \frac{V \cdot P \cdot F \cdot K_p}{K_d}, \quad (4.7)$$

Де:

V – вартість 1 кВт·год електроенергії, в 2025 р. $V \approx 4,5$ грн/кВт·год;

P – встановлена потужність обладнання, $P = 1,0$ кВт;

F – фактична кількість годин роботи обладнання, прийmemo $F = 310$ год;

K_p – коефіцієнт використання потужності ($K_p = 0,86$);

K_d – коефіцієнт корисної дії обладнання ($K_d = 0,79$). Тоді витрати на електроенергію становитимуть:

$$V_e = \frac{4,5 \times 1,0 \times 310 \times 0,86}{0,79} = 1518,5 \approx 1519 \text{ грн.}$$

4.2.8 Інші витрати

Інші витрати $V_{ін}$ можна прийняти як (50...300)% від основної заробітної плати розробників:

$$V_{ін} = (0,5 \dots 3,0) \times Z_0, \quad (4.8)$$

Для нашого випадку прийmemo коефіцієнт 0,90 (90%). Отже:

$$V_{ін} = 0,9 \times 37217 = 33495,3 \approx 33496 \text{ грн.}$$

4.2.9 Сума всіх попередніх статей витрат

Сума всіх попередніх статей витрат становить витрати на виконання роботи безпосередньо розробниками (магістрантом) – позначимо цю суму як V .

$$V = 37217 + 3722 + 9007 + 3454 + 1600 + 1519 + 33496 = 90015 \text{ грн.}$$

4.2.10 Загальні витрати на розробку методу автоматизованого генерування зображень

Загальні витрати на розробку методу автоматизованого генерування зображень можна спрогнозувати з урахуванням непередбачених витрат наступним виразом:

$$V_{заг} = V \cdot \left(1 + \frac{\beta}{100}\right), \quad (4.9)$$

де β – коефіцієнт, що враховує додаткові (непередбачені) витрати, %.

Приймаючи $\beta \approx 5\%$, одержимо:

$$V_{заг} \approx 90015 \times 1,05 \approx 94500 \text{ грн.}$$

Отже, прогнозовані загальні витрати на розробку методу автоматизованого генерування alt-тексту складають приблизно 95 тис. грн (з урахуванням округлення).

4.3 Теоретичний розрахунок економічного ефекту від можливої комерціалізації методу

У цьому підрозділі подано теоретичні підрахунки та аналітичні міркування щодо потенційної економічної ефективності розробленого методу автоматизованого генерування альтернативного тексту до зображень для вебресурсів. Станом на момент підготовки роботи розробка не впроваджена в комерційне використання, а оцінки базуються на гіпотетичних сценаріях можливого попиту та ринкової вартості.

Передбачається, що аналогічні за функціональністю інструменти на ринку коштують близько 95 тис. грн. Запропонований метод може бути конкурентним завдяки застосуванню сучасних моделей штучного інтелекту, мультиінструментальній архітектурі та високій швидкості генерації описів. У зв'язку з цим умовна ринкова ціна нашого методу оцінюється на рівні 115 тис. грн.

Прогнозна кількість потенційних споживачів, що можуть бути зацікавлені у впровадженні подібного рішення (розробники веб-сайтів, освітні та державні установи, digital-агенції), оцінюється на рівні близько 100 клієнтів у перші три роки. За оптимістичного сценарію очікується зростання попиту: +10 користувачів у першому році, +15 — у другому, +20 — у третьому.

Для кожного року було змодельовано можливе збільшення чистого прибутку інвестора за виразом:

$$\Delta\Pi_i = [(\Delta\Pi_0 \times N) + (\Pi_0 \times \Delta N_i)] \times \lambda \times \rho \times (1 - v), \quad (4.10)$$

де:

$\Delta\Pi_0 = 20$ тис. грн – приріст ціни;

$N = 100$ – базова кількість користувачів;

$\Delta N_i = 10, 15, 20$ – прогнозоване щорічне зростання;

$\Pi_0 = 115$ тис. грн – ціна одного впровадження;

$\lambda = 0,8333$ – коефіцієнт ПДВ;

$\rho = 0,45$ – коефіцієнт рентабельності;

$\nu = 0,18$ – ставка податку на прибуток.

На цій основі отримано прогнозні прибутки:

$\Delta\Pi_1 = 969$ тис. грн,

$\Delta\Pi_2 = 1146$ тис. грн,

$\Delta\Pi_3 = 1322$ тис. грн.

Розрахуємо приведену вартість (ПП) усіх чистих прибутків за 3 роки з використанням дисконтування за ставкою 10%:

$$\text{ПП} = \sum_{i=1}^3 \frac{\Delta\Pi_i}{(1+\tau)^i} = \frac{\Delta\Pi_1}{(1+\tau)^1} + \frac{\Delta\Pi_2}{(1+\tau)^2} + \frac{\Delta\Pi_3}{(1+\tau)^3}, \quad (4.11)$$

де $\tau = 0,10$.

Отримаємо: $\text{ПП} \approx 801 + 861 + 903 = 2\,565$ тис. грн.

Оцінимо теперішню вартість інвестицій: $PV = 5 \times 95 = 475$ тис. грн. Тоді абсолютний ефект:

$$E_{\text{абс}} = \text{ПП} - PV = 2565 - 475 = 2090 \text{ тис. грн.}$$

Розрахуємо внутрішню дохідність:

$$E_B = \left(\frac{E_{\text{абс}}}{PV} + 1 \right)^{\frac{1}{T_{\text{ж}}}} - 1, \quad (4.12)$$

де $T_{\text{ж}} = 4$ роки.

$$E_B = \left(\frac{2090}{475} + 1 \right)^{\frac{1}{4}} - 1 \approx 52,4\%$$

Оскільки $E_B > \tau_{\text{мін}} = 50\%$, інвестиційно розробка є привабливою в умовах

На основі проведених розрахунків визначимо основні техніко-економічні показники проєкту та порівняємо їх із запланованими (результати зведено у табл. 4.6).

Таблиця 4.6 – Порівняння техніко-економічних показників з очікуваними

Показник	Задано в ТЗ	Отримано в роботі	Висновок
Витрати на розробку	≈ 100 тис. грн	95 тис. грн	В межах запланованого
Абсолютний економічний ефект (розрахунковий)	≈ 2 млн грн	2,09 млн грн (теоретично)	Може бути досягнутий за сприятливих умов
Внутрішня дохідність інвестицій (розрахункова)	≥ 50%	52,4% (модельна оцінка)	Потенційно досяжна
Орієнтовний термін окупності	≤ 3 роки	1,9 роки (за моделлю)	Теоретично в межах очікуваного

Таким чином, основні техніко-економічні показники розробленого нами методу автоматизованого генерування alt-тексту, визначені у технічному завданні, виконані. Отримані результати підтверджують доцільність комерціалізації методики за умови сприятливих макроекономічних показників (помірної інфляції) та дозволяють рекомендувати впровадження розробки в практику.

4.4 Висновки до розділу

У четвертому розділі проведено техніко-економічне обґрунтування розробки та впровадження методу автоматизованого генерування альтернативного тексту для зображень. Визначено структуру витрат, спрогнозовано потенційний прибуток та оцінено інвестиційну привабливість проєкту.

Аналіз витрат. Загальна оціночна вартість розробки та виведення продукту на ринок (або вартість аналога) становить 95 000 грн. Ця сума включає витрати на оплату праці розробників, використання серверних потужностей, маркетингові заходи та інші накладні витрати.

Економічний ефект. Розрахунки показали високу потенційну ефективність впровадження методу. Абсолютний економічний ефект (чистий дисконтований дохід за 3 роки) прогнозується на рівні 2,09 млн грн. Це досягається за рахунок автоматизації рутинних процесів, які при ручному виконанні потребують значних часових та фінансових ресурсів.

Інвестиційна привабливість. Розрахункова внутрішня норма дохідності (Ев) становить 52,4%, що перевищує встановлене порогове значення (50%). Це свідчить про доцільність інвестування у проєкт за умов стабільної економічної ситуації та помірної інфляції (до 10%).

Окупність. Орієнтовний термін окупності проєкту становить 1,9 року, що вписується у заплановані рамки (до 3 років) і підтверджує ліквідність розробки.

Ризики. Моделювання впливу інфляції показало, що при зростанні рівня інфляції до 20–30% показники дохідності знижуються до 42,8% та 34,6% відповідно. Це вказує на необхідність розробки додаткових заходів мінімізації ризиків при виході на ринок.

ВИСНОВКИ

У магістерській кваліфікаційній роботі розв'язано актуальну науково-прикладну задачу підвищення рівня вебдоступності цифрових ресурсів шляхом автоматизації формування альтернативного тексту для зображень. Отримані результати дозволяють зробити такі висновки.

Проведений аналіз сучасного стану вебдоступності показав, що відсутність або некоректність ALT-текстів є однією з найбільш поширених причин недоступності вебресурсів для користувачів із порушеннями зору. Встановлено, що навіть за наявності нормативних вимог значна частина державних, освітніх та корпоративних сайтів не відповідає базовим критеріям доступності.

Дослідження міжнародних і національних нормативних документів підтвердило, що вимога щодо надання текстових альтернатив для зображень є обов'язковою та юридично значущою. Виконання стандартів WCAG 2.1 та ДСТУ EN 301 549:2022 розглядається не як рекомендація, а як необхідна умова функціонування сучасних цифрових сервісів.

Аналіз існуючих підходів до формування ALT-текстів засвідчив, що ручний метод забезпечує найвищу якість, але є непридатним для масштабних вебресурсів через високу трудомісткість. Автоматизовані підходи, навпаки, мають високу продуктивність, однак потребують адаптації та контролю для відповідності вимогам доступності.

Обґрунтовано доцільність використання мультимодальних моделей штучного інтелекту для задачі автоматизованого генерування ALT-текстів за умови їх спеціальної адаптації. Показано, що без додаткових обмежень такі моделі схильні генерувати описові або суб'єктивні тексти, які не відповідають функціональному призначенню ALT-описів.

Запропоновано метод автоматизованого генерування альтернативного тексту на основі мультимодальної моделі BLIP-2, який орієнтований на створення коротких, нейтральних і семантично коректних описів відповідно

до вимог стандартів вебдоступності. Метод враховує особливості використання зображень у вебконтенті та їхню роль у структурі сторінки.

Розроблено архітектуру програмного модуля автоматичного генерування ALT-текстів із локальним серверним API та механізмом інтеграції з системою управління контентом Joomla. Обране архітектурне рішення забезпечує автономність роботи, збереження конфіденційності даних та можливість використання у державних і освітніх установах.

Реалізований програмний модуль продемонстрував працездатність і практичну доцільність використання в реальних умовах. Результати тестування підтвердили можливість автоматизованого формування альтернативних текстів прийнятної якості та зменшення навантаження на контент-редакторів.

Практична цінність роботи полягає у можливості впровадження розробленого рішення для підвищення рівня цифрової інклюзії, забезпечення відповідності вебресурсів чинним стандартам доступності та підтримки державної політики безбар'єрності.

Перспективи подальших досліджень полягають у розширенні підтримки інших систем управління контентом, удосконаленні механізмів контекстного аналізу ролі зображень, а також у використанні гібридних підходів із залученням експертної верифікації згенерованих ALT-текстів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Національна стратегія зі створення безбар'єрного простору в Україні до 2030 року. URL: <https://zakon.rada.gov.ua/laws/show/366-2021-%D1%80> (дата звернення 03.12.2025)
2. Закон України «Про основні засади забезпечення доступності вебресурсів органів влади». URL: <https://zakon.rada.gov.ua/laws/show/2773-20> (дата звернення 03.12.2025)
3. W3C Web Accessibility Initiative (WAI). URL: <https://www.w3.org/WAI/>
4. WHO Vision Impairment Global Report. URL: <https://www.who.int/publications/i/item/world-report-on-vision> (дата звернення 03.12.2025)
5. Міністерство соціальної політики України — дані про інвалідність. URL: <https://www.msp.gov.ua/news/12292.html> (дата звернення 03.12.2025)
6. European Commission Digital Accessibility Findings. URL: <https://ec.europa.eu/digital-strategy/our-policies/web-accessibility> (дата звернення 03.12.2025)
7. WebAIM Million Report 2023. URL: <https://webaim.org/projects/million/> (дата звернення 03.12.2025)
8. WCAG 2.1 Standard Documentation. URL: <https://www.w3.org/TR/WCAG21/> (дата звернення 03.12.2025)
9. European Accessibility Act (Directive 2019/882/EU). URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32019L0882> (дата звернення 03.12.2025)
10. Multimodal Model Hallucination Review (2023). URL: <https://arxiv.org/abs/2311.00059> (дата звернення 03.12.2025)
11. European Commission. ICT Accessibility Overview. URL: <https://ec.europa.eu/social/main.jsp?catId=1202> (дата звернення 03.12.2025)

12. W3C. Web Content Accessibility Guidelines (WCAG) 2.1. URL: <https://www.w3.org/TR/WCAG21/> (дата звернення 03.12.2025)
13. ДСТУ EN 301 549:2022 — Національний стандарт доступності. URL: <https://uas.gov.ua/standard/dstu-en-301-549-2022> (дата звернення 03.12.2025)
14. U.S. Department of Justice. ADA Compliance Reports 2023. URL: <https://www.ada.gov/resources/web-guidance/> (дата звернення 03.12.2025)
15. WebAIM Million Report 2023 — Accessibility Analysis. URL: <https://webaim.org/projects/million/> (дата звернення 03.12.2025)
16. W3C WAI — ALT Text Techniques. URL: <https://www.w3.org/WAI/tutorials/images/decision-tree/> (дата звернення 03.12.2025)
17. Nielsen Norman Group — UX Impact of Alt Attributes. URL: <https://www.nngroup.com/articles/image-alt-text/> (дата звернення 03.12.2025)
18. Accessibility Foundation Review — Common ALT Mistakes. URL: <https://accessibility.site/images/alt-text.html> (дата звернення 03.12.2025)
19. Nielsen Norman Group. Accessibility and User Experience Study. URL: <https://www.nngroup.com/topic/accessibility/> (дата звернення 03.12.2025)
20. Google Search Central: Image SEO Guidelines. URL: <https://developers.google.com/search/docs/appearance/google-images> (дата звернення 03.12.2025)
21. W3C Techniques for ALT Text in Educational Materials. URL: <https://www.w3.org/WAI/WCAG21/Techniques/general/G94> (дата звернення 03.12.2025)
22. Accessibility Guidelines for Higher Education Resources. URL: <https://www.washington.edu/accessibility/web/images/> (дата звернення 03.12.2025)
23. W3C Authoring Techniques for WCAG 2.1. URL: <https://www.w3.org/WAI/WCAG21/quickref/> (дата звернення 03.12.2025)
24. Web Accessibility Initiative — Best Practices for ALT. URL: <https://www.w3.org/WAI/tips/images/> (дата звернення 03.12.2025)

25. Microsoft Cognitive Services Captioning Research. URL: <https://azure.microsoft.com/en-us/products/ai-services/ai-vision> (дата звернення 03.12.2025)
26. Google AI Multimodal Limitations Study. URL: <https://ai.google/discover/research/> (дата звернення 03.12.2025)
27. Meta AI: Challenges in Image Captioning Systems. URL: <https://ai.meta.com/research/> (дата звернення 03.12.2025)
28. Google Research: Hallucination Study in Vision-Language Models. URL: <https://arxiv.org/abs/2311.00059> (дата звернення 03.12.2025)
29. Meta AI: Multimodal Model Limitations. URL: <https://ai.meta.com/blog/> (дата звернення 03.12.2025)
30. Microsoft Research: Captioning vs ALT Text. URL: <https://www.microsoft.com/en-us/research/publication/captioning-images-for-accessibility/> (дата звернення 03.12.2025)
31. European Data Security Standards 2023. URL: https://edpb.europa.eu/our-work-tools/general-guidance/gdpr-guidelines-recommendations-best-practices_en (дата звернення 03.12.2025)
32. W3C: Cognitive and Ethical Risks in Automated Accessibility Tools. URL: <https://www.w3.org/TR/naa/> (дата звернення 03.12.2025)
33. HTTP Archive. Web Almanac 2023 — Image Analysis. URL: <https://almanac.httparchive.org/en/2022/media> (дата звернення 03.12.2025)
34. NVDA User Guide. Image Accessibility. URL: <https://www.nvaccess.org/>
35. Joomla Documentation. Media Manager & Accessibility. URL: <https://docs.joomla.org/> (дата звернення 03.12.2025)
36. Research on Accessibility of Ukrainian Universities Web Resources (2023). URL: <https://mon.gov.ua/> (дата звернення 03.12.2025)
37. Li J. BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models. URL: <https://arxiv.org/abs/2301.12597> (дата звернення 03.12.2025)

38. OpenVINO. CPU Inference Benchmarking for Vision-Language Models. URL: <https://docs.openvino.ai/> (дата звернення 03.12.2025)
39. FastAPI Documentation. High Performance Python Web Framework. URL: <https://fastapi.tiangolo.com/> (дата звернення 03.12.2025)
40. SQLite Documentation. Hash-Based Caching Mechanisms. URL: <https://sqlite.org/docs.html> (дата звернення 03.12.2025)
41. Козловський В.О., Лесько О.Й., Кавецький В.В. Методичні вказівки до виконання економічної частини магістерських кваліфікаційних робіт. – Вінниця: ВНТУ, 2021. – 42 с. URL: <https://library.vntu.edu.ua/e-books/ekonomichna-chastyna-mkr-2021.pdf> (дата звернення 03.12.2025)
42. BradyFU. Awesome Multimodal Large Language Models: Evaluation. GitHub repository, 2024. URL: <https://github.com/BradyFU/Awesome-Multimodal-Large-Language-Models/tree/Evaluation> (дата звернення 03.12.2025)
43. МЕТОД АВТОМАТИЗОВАНОГО ГЕНЕРУВАННЯ АЛЬТЕРНАТИВНОГО ТЕКСТУ ДЛЯ ЗОБРАЖЕНЬ У ВЕБРЕСУРСАХ / Д.А.Кавунець, І.В.Богач// Матеріали міжнародної науково-практичної Інтернет-конференції студентів, аспірантів та молодих науковців «Молодь в науці: дослідження, проблеми, перспективи (МН-2025-26)». Збірник доповідей [Електронний ресурс] Режим доступу: <https://conferences.vntu.edu.ua/index.php/mn/mn2026/paper/view/26818>

ДОДАТКИ

Додаток А (обов'язковий)

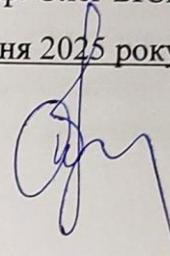
Технічне завдання

ЗАТВЕРДЖУЮ

Завідувач кафедри АІТ

д.т.н., проф. Олег БІСІКАЛО

«17» жовтня 2025 року



ТЕХНІЧНЕ ЗАВДАННЯ

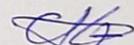
на магістерську кваліфікаційну роботу

**«МЕТОД АВТОМАТИЗОВАНОГО ГЕНЕРУВАННЯ АЛЬТЕРНАТИВНОГО
ТЕКСТУ ДЛЯ ЗОБРАЖЕНЬ У ВЕБРЕСУРСАХ»**

08-31.МКР.004.02.000 ТЗ

Керівник роботи:

к.т.н., доц. каф. АІТ

Ярослав КУЛИК 

«16» жовтня 2025 р.

Виконавець:

ст. гр. 1АКІТР-24м

Дар'я КАВУНЕЦЬ 

«16» жовтня 2025 р.

Вінниця ВНТУ – 2025

1. Назва та галузь застосування

Метод автоматизованого генерування альтернативного тексту для зображень у вебресурсах.

Інформаційні технології, вебдоступність (Web Accessibility), цифрова інклюзія, системи комп'ютерного зору та штучного інтелекту.

2. Підстава для розробки

Розробку системи здійснювати на підставі наказу по університету № 313 від 24 вересня 2025 року та завдання до магістерської кваліфікаційної роботи, складеного та затвердженого кафедрою «Автоматизації та інтелектуальних інформаційних технологій»

3. Мета та призначення розробки

Метою роботи є розробка методу автоматизованого формування альтернативного тексту до зображень у вебресурсах, що забезпечує:

- Автоматичне розпізнавання змісту зображення з використанням мультимодальної моделі BLIP-2;
- Формування короткого та змістовного ALT-тексту українською мовою відповідно до вимог WCAG 2.1;
- Нормалізацію тексту та видалення галюцинацій моделі;
- Автоматичну інтеграцію сервісу з CMS Joomla для виявлення зображень без ALT-атрибутів;
- Кешування результатів генерації для оптимізації продуктивності;
- Можливість REST-викликів для інтеграції з іншими вебресурсами.

Призначення системи: забезпечення коректної вебдоступності, зменшення трудовитрат редакторів контенту та підвищення відповідності сайту стандартам цифрової інклюзії.

4. Джерела розробки

1. Web Content Accessibility Guidelines (WCAG) 2.1. URL: <https://www.w3.org/TR/WCAG21/>

2. Kabachii V. Identifying moments of decision making on trade in financial time series using fuzzy cluster analysis / V. Kabachii, R. Maslii, S. Kozlovskiy, O. Dronchack // Neuro-Fuzzy Modeling Techniques in Economics. – 2023. – № 12. – P. 175–192. – DOI: 10.33111/nfmte.2023.175.

3. ДСТУ EN 301 549:2022. Інформаційні технології. Вимоги щодо доступності продуктів та послуг ІКТ.

4. FastAPI Documentation. URL: <https://fastapi.tiangolo.com/>

5. Salesforce BLIP-2 Model (LAVIS). URL: <https://github.com/salesforce/LAVIS>

5. Показники призначення

5.1. Основні технічні характеристики системи

Функціональні можливості:

5.1.1 Модуль обробки зображень та генерації ALT-тексту:

- Завантаження вхідних зображень у форматах: JPEG, PNG, WEBP;
- Стандартизація та нормалізація зображення перед подачею в модель;
- Використання мультимодальної моделі BLIP-2 для генерації початкового текстового опису;
- Використання додаткових фільтрів для усунення галюцинацій та надмірної деталізації;
- Формування фінального ALT-тексту довжиною 5–30 слів.

5.1.2 Модуль REST-API сервісу:

- Отримання зображення через API з поверненням ALT-опису;
- Підтримка JSON-відповідей та кодувань UTF-8;
- Обробка пакетних запитів.

5.1.3 Модуль інтеграції з CMS Joomla:

- Автоматичне виявлення зображень без ALT-атрибутів;
- Передача файлу до API-сервісу та отримання опису;
- Запис альтернативного тексту в адміністративну панель Joomla.

5.1.4 Модуль кешування:

- Збереження результатів генерації в локальній базі SQLite;
- Перевірка наявності попередньо згенерованих записів для уникнення повторної обробки.

7. Стадії розробки:

1. Розділ 1 «Аналіз предметної області та існуючих методів» має бути виконаний до 05.10.2025 р.
2. Розділ 2 «Методи та засоби автоматизованого генерування ALT-тексту» має бути виконаний до 25.10.2025 р.
3. Розділ 3 «Розробка програмного забезпечення для формування ALT-тексту» має бути виконаний до 20.11.2025 р.
4. Розділ 4 «Економічний розділ» має бути виконаний до 01.12.2025 р.

8. Порядок контролю та приймання

1. Рубіжний контроль провести до 14.11.2025.
2. Попередній захист магістерської кваліфікаційної роботи провести до 02.12.2025.
3. Захист магістерської кваліфікаційної роботи провести в період з 15.12.2025 р. до 19.12.2025 р.

Додаток Б
(обов'язковий)

ІЛЮСТРАТИВНА ЧАСТИНА

МЕТОД АВТОМАТИЗОВАНОГО ГЕНЕРУВАННЯ АЛЬТЕРНАТИВНОГО
ТЕКСТУ ДЛЯ ЗОБРАЖЕНЬ У ВЕБРЕСУРСАХ

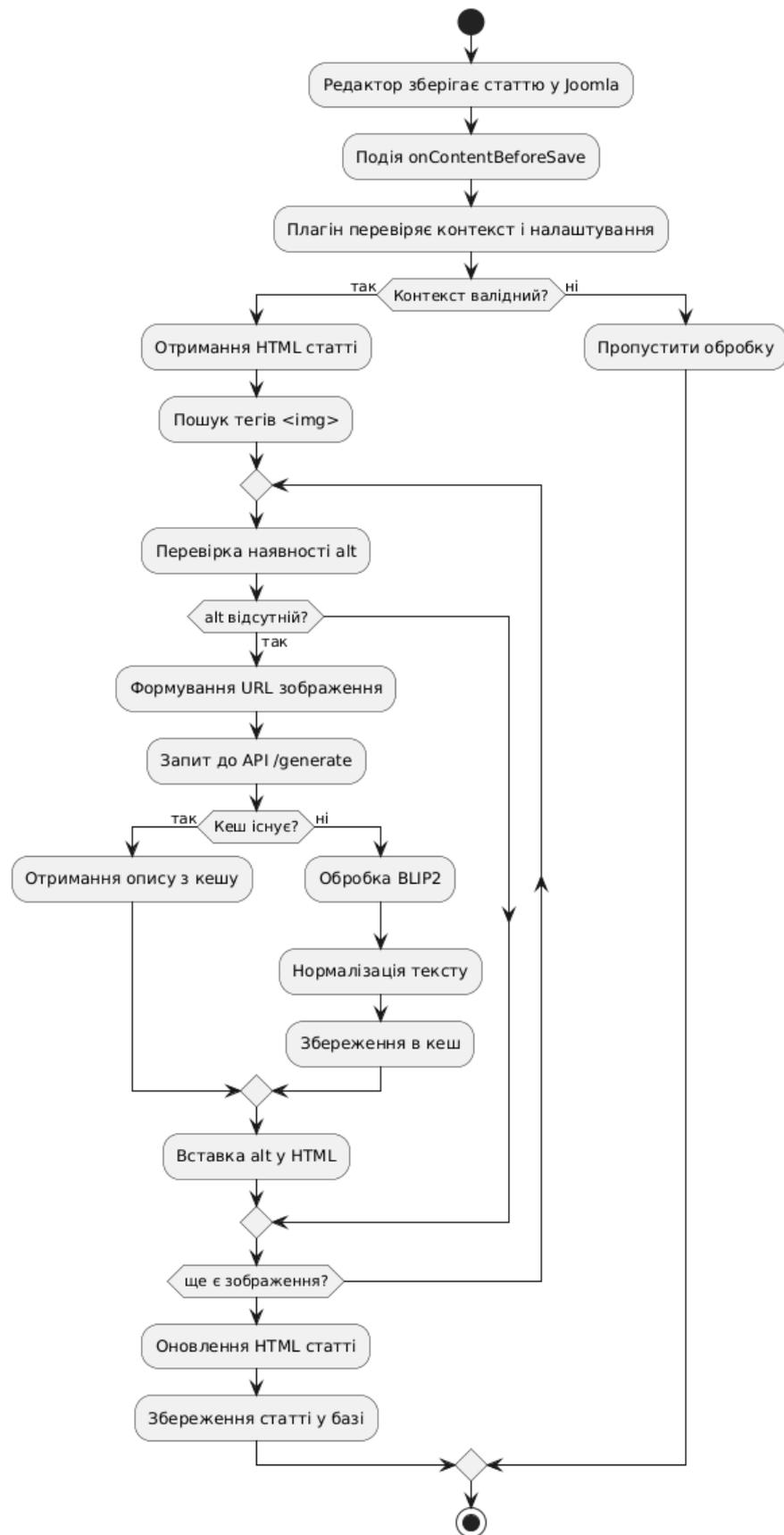


Рисунок Б.1 — Архітектура програмного модуля автоматизованого генерування ALT-текстів (плагін Joomla та API-сервіс)

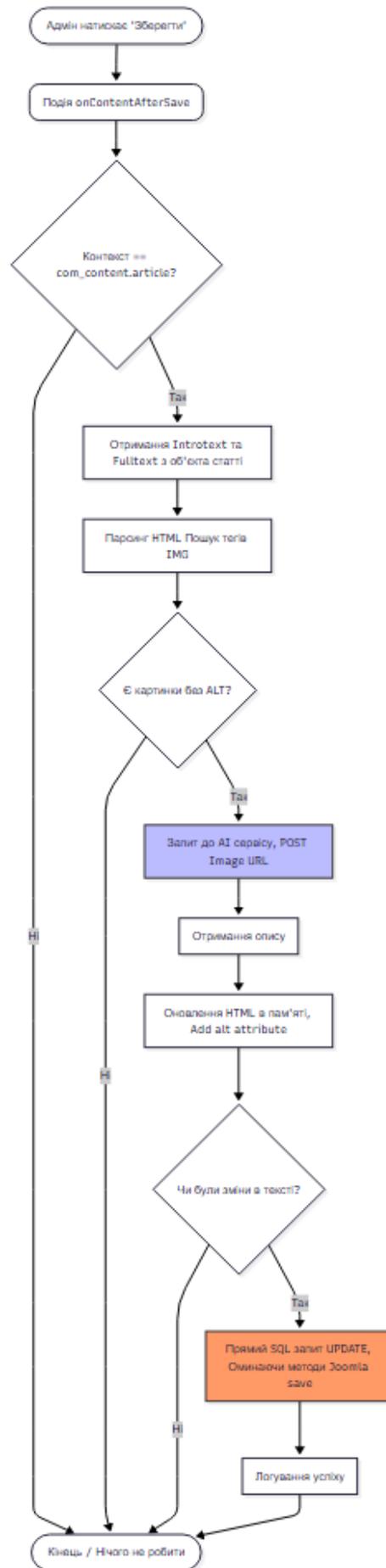


Рисунок Б.2 — Структура плагіну автоматичного генерування alt-тексту

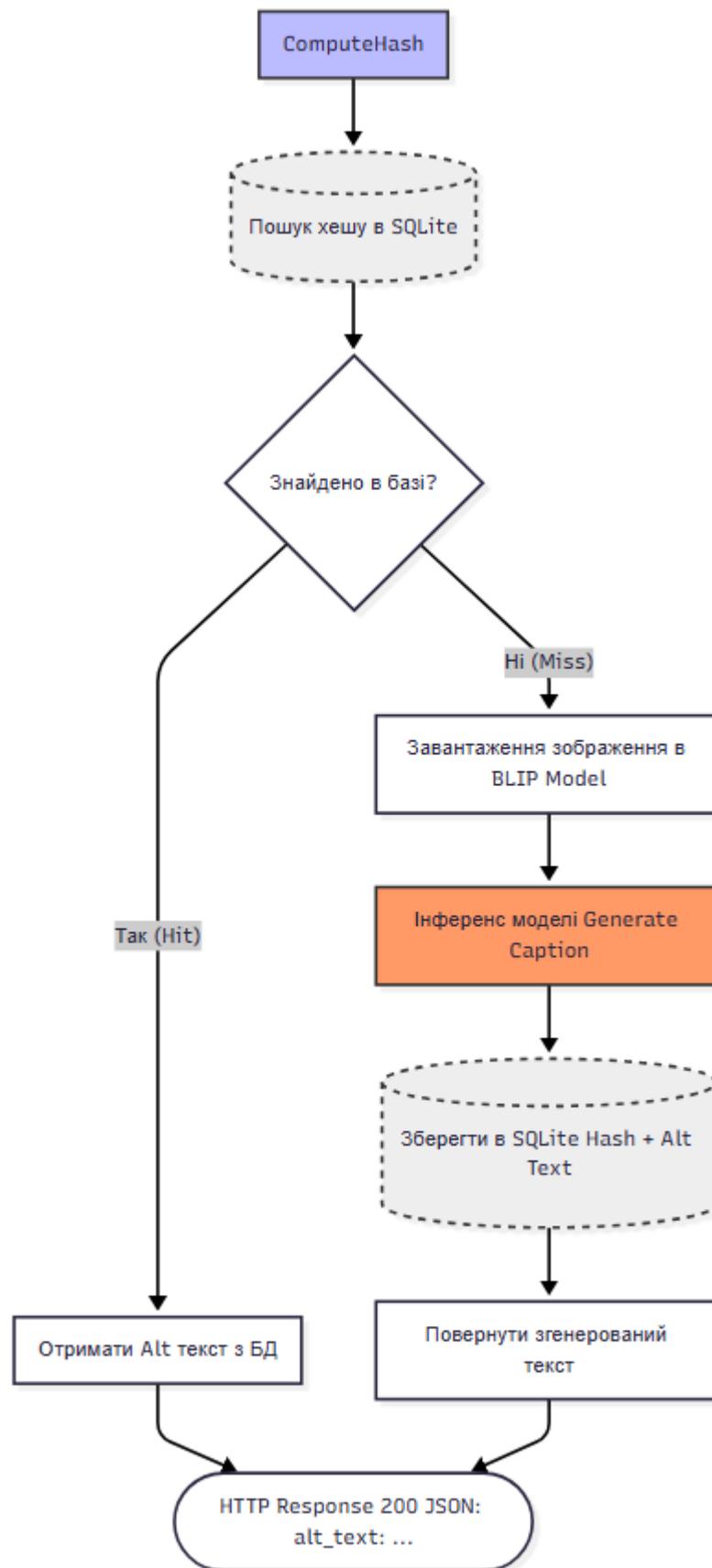


Рисунок Б.3 — Приклад роботи REST-API сервісу

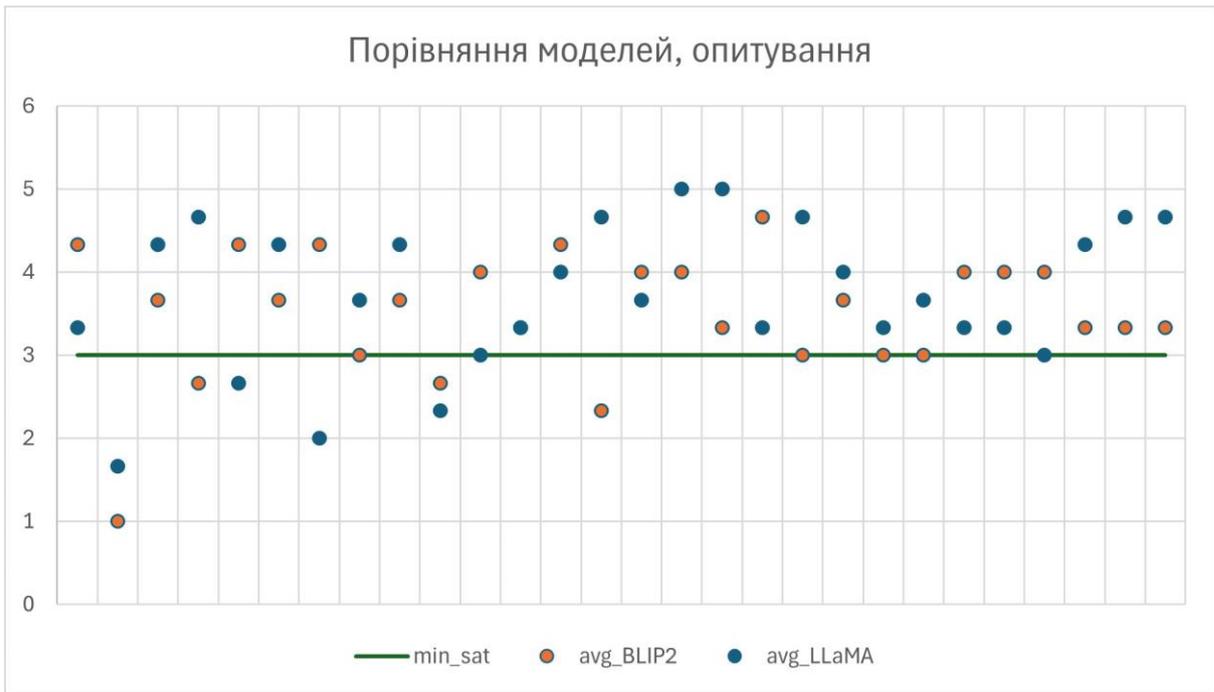


Рисунок Б.4 — Порівняльна точкова діаграма оцінювання моделей

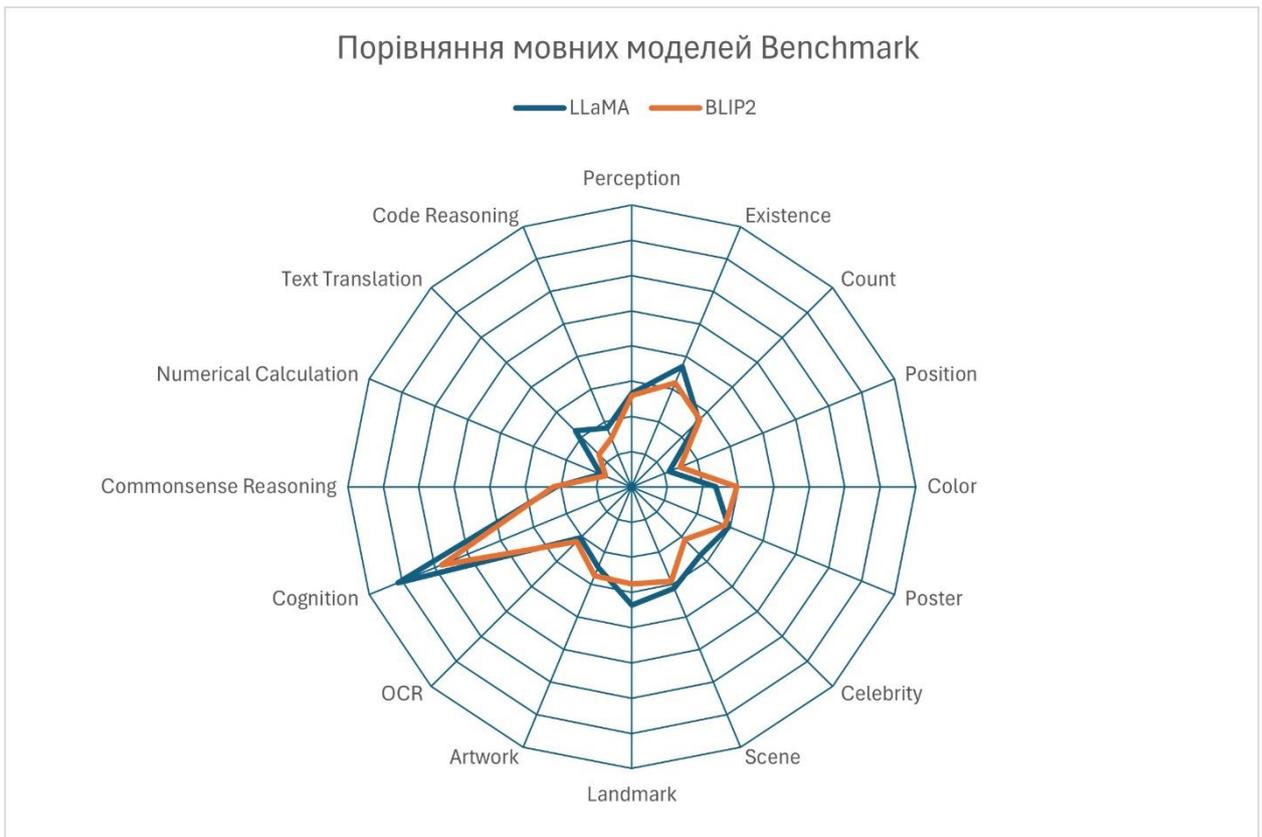


Рисунок Б.5 — Порівняння моделей за категоріями Multimodal Benchmark

```

Командний рядок
C:\Users\user>curl -X POST "http://localhost:8000/analyze" -H "Content-Type: application/json" -d '{"image_url": "http://localhost:8000/img/photo1.jpeg", "lang": "uk"}'
{"alt_text": "за столом сидять чоловік і жінка з ноутбуком", "source": "cache"}
C:\Users\user>curl -X POST "http://localhost:8000/analyze" -H "Content-Type: application/json" -d '{"image_url": "http://localhost:8000/img/photo5.jpeg", "lang": "uk"}'
{"alt_text": "двоє чоловіків сидять за столом з комп'ютером", "source": "generated"}
C:\Users\user>

```

Рисунок Б.6 – Перевірка роботи API-сервісу через командний рядок



Рисунок Б.7 – Компоненти принципу «Сприйняття» стандарту WCAG 2.1

Принципи доступності

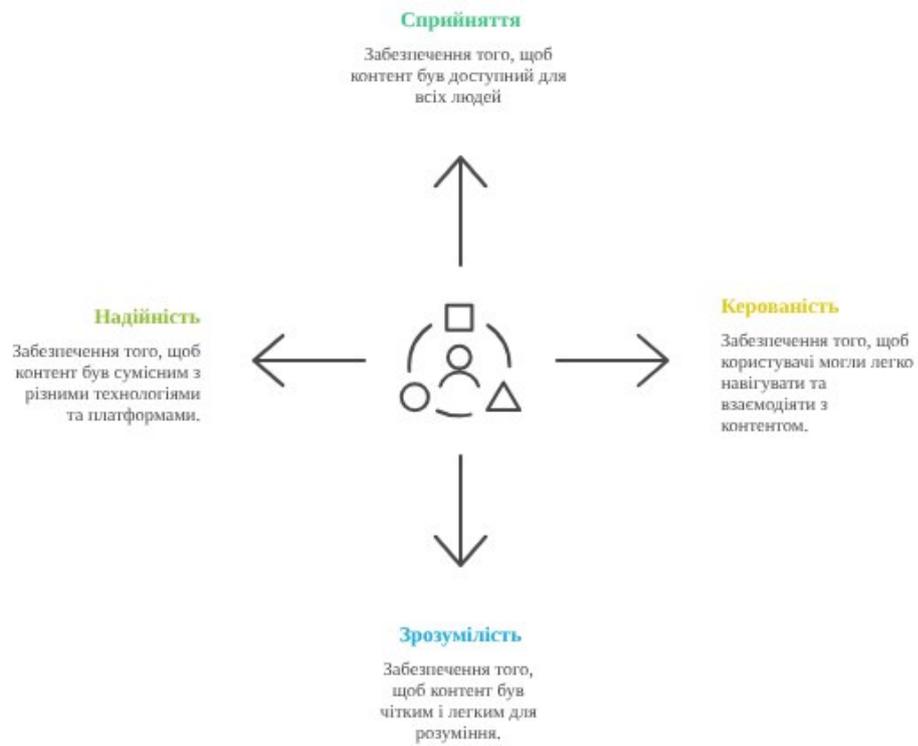


Рисунок Б.8 – Основні принципи доступності стандарту WCAG 2.1

Додаток В

Лістинг основних функцій програмного забезпечення

```
import sqlite3
import hashlib
import io
import requests
import torch
import uvicorn

from fastapi import FastAPI, HTTPException
from pydantic import BaseModel
from PIL import Image
from transformers import BlipProcessor, BlipForConditionalGeneration
from deep_translator import GoogleTranslator

#config
MODEL_ID = "Salesforce/blip-image-captioning-large"
DB_FILE = "image_captions.db"
DEVICE = "cuda" if torch.cuda.is_available() else "cpu"

app = FastAPI(title="Image Alt Generator Service")

#модель
print(f"Loading model {MODEL_ID} on {DEVICE}...")
processor = BlipProcessor.from_pretrained(MODEL_ID)
model =
BlipForConditionalGeneration.from_pretrained(MODEL_ID).to(DEVICE)
print("Model loaded successfully!")

#База даних
def init_db():
```

```

with sqlite3.connect(DB_FILE) as conn:
    conn.execute("""
        CREATE TABLE IF NOT EXISTS captions (
            hash TEXT PRIMARY KEY,
            alt_text_en TEXT,
            alt_text_uk TEXT
        )
    """)
init_db()

class ImageRequest(BaseModel):
    image_url: str
    lang: str = "uk" # uk = українська, en = англійська

def get_image_hash(image_bytes: bytes) -> str:
    """Генерує SHA-256 хеш для байтів зображення"""
    return hashlib.sha256(image_bytes).hexdigest()

def translate_text(text: str, target_lang: str = 'uk') -> str:
    """Перекладає текст (BLIP видає англійську)"""
    try:
        return GoogleTranslator(source='auto', target=target_lang).translate(text)
    except Exception as e:
        print(f"Translation error: {e}")
        return text

#endpoint
@app.post("/analyze")
def analyze_image(request: ImageRequest):
    try:

```

```

# 1. Завантажуємо зображення за URL
response = requests.get(request.image_url, stream=True, timeout=10)
response.raise_for_status()
image_bytes = response.content

# 2. Обчислюємо хеш
img_hash = get_image_hash(image_bytes)

# 3. Перевіряємо кеш (SQLite)
with sqlite3.connect(DB_FILE) as conn:
    cursor = conn.cursor()
    cursor.execute("SELECT alt_text_en, alt_text_uk FROM captions WHERE
hash = ?", (img_hash,))
    row = cursor.fetchone()

    if row:
        print(f"Cache HIT for {img_hash[:8]}...")
        return {
            "alt_text": row[1] if request.lang == 'uk' else row[0],
            "source": "cache"
        }

# 4. Якщо в кеші немає - генеруємо (Cache MISS)
print(f"Cache MISS for {img_hash[:8]}... Generating...")

raw_image = Image.open(io.BytesIO(image_bytes)).convert('RGB')

# Підготовка вхідних даних для моделі
inputs = processor(raw_image, return_tensors="pt").to(DEVICE)

```

```

# Генерація опису
out = model.generate(**inputs, max_new_tokens=50)
caption_en = processor.decode(out[0], skip_special_tokens=True)

# 5. Переклад українською
caption_uk = translate_text(caption_en, 'uk')

# 6. Зберігаємо результат в БД
with sqlite3.connect(DB_FILE) as conn:
    conn.execute(
        "INSERT INTO captions (hash, alt_text_en, alt_text_uk) VALUES (?, ?,
?)",
        (img_hash, caption_en, caption_uk)
    )

return {
    "alt_text": caption_uk if request.lang == 'uk' else caption_en,
    "source": "generated"
}

except requests.exceptions.RequestException as e:
    raise HTTPException(status_code=400, detail=f"Failed to fetch image:
{str(e)}")

except Exception as e:
    print(e)
    raise HTTPException(status_code=500, detail=str(e))

if __name__ == "__main__":
    # Запуск сервера
    uvicorn.run(app, host="0.0.0.0", port=8000)

```

Додаток Г



ВІННИЦЬКА ОБЛАСНА ВІЙСЬКОВА АДМІНІСТРАЦІЯ
 вул. Соборна, 70, м. Вінниця, 21050, тел. (0432) 50-42-00,
 e-mail: oda@vin.gov.ua, web: www.vin.gov.ua
 Код ЄДРПОУ 20089290

09.12.2025 № 01.01-57/10567

На № _____ від _____

АКТ

впровадження результатів магістерської кваліфікаційної роботи

Кавунець Дар'ї Андріївни «Метод автоматизованого генерування альтернативного тексту для зображень у вебресурсах»

Комісія у складі керівника апарату обласної військової адміністрації О. Коханця, начальника управління цифрової трансформації апарату обласної військової адміністрації Л. Попихач, директора Департаменту інформаційної діяльності та комунікацій з громадськістю обласної військової адміністрації С. Юзвак склали цей акт про те, що у Вінницькій обласній військовій адміністрації впроваджуються результати, що отримані магістром Кавунець Д.А. при виконанні магістерської кваліфікаційної роботи, мають практичну цінність.

Розроблений метод автоматичного генерування альтернативного тексту для зображень за допомогою технологій штучного інтелекту дозволяє значно підвищити доступність веб-ресурсів, скоротити навантаження на контент-менеджерів та забезпечити відповідність державних інформаційних ресурсів міжнародним стандартам доступності. Запропоноване рішення вже проходить тестування на реальних даних і може бути інтегроване в існуючі системи управління контентом органу влади, що підтверджує його практичну придатність та економічну ефективність.

Вінницькій обласній військовій адміністрації магістром Кавунець Д.А. передано програмне забезпечення для автоматизованого генерування альтернативного тексту до зображень, яке впроваджується в органі влади з метою підвищення їх інформаційної та доступної якості.

Члени комісії:

Керівник апарату обласної
військової адміністрації

Начальник управління цифрової
трансформації апарату обласної
військової адміністрації

Директор Департаменту
інформаційної діяльності та
комунікацій з громадськістю
обласної військової адміністрації



Олександр КОХАНЕЦЬ

Людмила ПОПИХАЧ

Світлана ЮЗВАК

Додаток Д (обов'язковий)

ПРОТОКОЛ ПЕРЕВІРКИ КВАЛІФІКАЦІЙНОЇ РОБОТИ

Назва роботи: «Метод автоматизованого генерування альтернативного тексту для зображень у вебресурсах»

Тип роботи: магістерська кваліфікаційна робота
(бакалаврська кваліфікаційна робота / магістерська кваліфікаційна робота)

Підрозділ кафедра АІТ
(кафедра, факультет, навчальна група)

Коефіцієнт подібності текстових запозичень, виявлених у роботі системою StrikePlagiarism (КПІ) 1.18 %

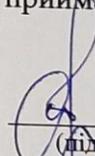
Висновок щодо перевірки кваліфікаційної роботи (відмітити потрібне)

- Запозичення, виявлені у роботі, оформлені коректно і не містять ознак академічного плагіату, фабрикації, фальсифікації. Роботу прийняти до захисту.
- У роботі не виявлено ознак плагіату, фабрикації, фальсифікації, але надмірна кількість текстових запозичень та/або наявність типових розрахунків не дозволяють прийняти рішення про оригінальність та самостійність її виконання. Роботу направити на доопрацювання.
- У роботі виявлено ознаки академічного плагіату та/або в ній містяться навмисні спотворення тексту, що вказують на спроби приховування недобросовісних запозичень. Робота до захисту не приймається.

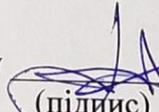
Експертна комісія:

Бісікало О.В., зав. каф. АІТ
(прізвище, ініціали, посада)

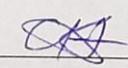
Овчинников К.В., доц. каф. АІТ
(прізвище, ініціали, посада)


(підпис)

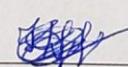
(підпис)

Особа, відповідальна за перевірку  Маслій Р.В.
(підпис) (прізвище, ініціали)

З висновком експертної комісії ознайомлений(-на)

Керівник 
(підпис)

Кулик Я. А., доц. каф. АІТ
(прізвище, ініціали, посада)

Студент 
(підпис)

Кавунець Д. А.
(прізвище, ініціали)