

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Західноукраїнський національний університет
Факультет комп'ютерних інформаційних технологій
Кафедра інформаційно-обчислювальних систем і управління

КЛИМЧУК Анастасія Богданівна

**Інтерактивний застосунок для вивчення англійської мови на основі
медіа-контенту із використанням OpenAI / An interactive application
for learning English based on media-content using OpenAI**

Спеціальність 122 – Комп'ютерні науки
Освітньо-професійна програма – Комп'ютерні науки

Кваліфікаційна робота

Виконала студентка групи КН-42
А.Б. Климчук

Науковий керівник:
к.т.н., доцент І. В. Турченко

Кваліфікаційну роботу допущено до
захисту
«___»_____ 2025 р.

В.о. завідувача кафедри
_____ Н. М. Васильків

Тернопіль – 2025

Факультет комп'ютерних інформаційних технологій
Кафедра інформаційно-обчислювальних систем і управління
Освітній ступінь «бакалавр»
спеціальність 122 – Комп'ютерні науки
освітньо-професійна програма – Комп'ютерні науки

ЗАТВЕРДЖУЮ
В.о. завідувача кафедри
Н.М. Васильків
« ____ » _____ 2024 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТЦІ
КЛИМЧУК Анастасії Богданівні
(прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи: Інтерактивний застосунок для вивчення англійської мови на основі медіа-контенту із використанням OpenAI / An Interactive Application for Learning English Based on Media Content Using OpenAI

керівник роботи Турченко І. В., к.т.н., доцент

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом по університету від 20 грудня 2024 р. № 938.

2. Строк подання студентом закінченої кваліфікаційної роботи 25 травня 2025 р.

3. Вихідні дані до кваліфікаційної роботи: завдання на кваліфікаційну роботу студента, наукові статті, технічна література.

4. Основні питання, які потрібно розробити:

- аналіз предметної області;
- огляд існуючих аналогів і постановка задачі проектування;
- опис структури проєктованого інтерактивного застосунку;
- розробка алгоритмічного та інформаційного забезпечення;
- розробка інтерфейсу та функціоналу застосунку, що включає можливість перегляду медіа з інтерактивними елементами для вивчення мови;
- інтеграція дієвих інструментів вивчення мови, таких як субтитри, повторення, переклад, визначення нових слів та виразів в контексті;
- тестування різних форм інтерактивності;
- тестування застосунку та визначення можливих обмежень і недоліків при використанні.

5. Перелік графічного матеріалу в роботі:

- схема алгоритму роботи із субтитрами;
- схема алгоритму роботи із бібліотекою медіа-контенту;
- схема алгоритму інтерактивної взаємодії із субтитрами;
- схема алгоритму роботи із бібліотекою лексики;
- схема алгоритму вивчення лексики;
- схема алгоритму модуля роботи з медіа-контентом;
- схема алгоритму модуля роботи із лексикою;
- схема алгоритму модуля аналітики;
- схема алгоритму модуля управління резервними даними.

6. Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		Завдання видав	Завдання прийняв

7. Дата видачі завдання 20 грудня 2024 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів кваліфікаційної роботи	Примітка
1	Затвердження теми кваліфікаційної роботи, ознайомлення з літературними джерелами та складання плану роботи	до 01.01. 2025 р.	
2	Написання 1 розділу кваліфікаційної роботи	до 01.02. 2025 р.	
3	Написання 2 розділу кваліфікаційної роботи	до 01.04.2025 р.	
4	Написання 3 розділу кваліфікаційної роботи	до 01.05. 2025 р.	
5	Представлення попереднього варіанту кваліфікаційної роботи, перевірка та внесення змін керівником	до 15.05.2025 р.	
6	Опрацювання зауважень та представлення завершеного варіанту кваліфікаційної роботи. Підготовка супроводжуючих документів	до 25.05.2025 р.	
7	Перевірка кваліфікаційної роботи на оригінальність тексту	до 30.05.2025 р.	
8	Оформлення кваліфікаційної роботи та отримання допуску до захисту	до 10.06.2025 р.	
9	Подання кваліфікаційної роботи до захисту на засіданні атестаційної комісії	до 10.06. 2025 р.	

Студент _____ Климчук А.Б.
(підпис) (прізвище та ініціали)

Керівник кваліфікаційної роботи _____ Турченко І.В.
(підпис) (прізвище та ініціали)

АНОТАЦІЯ

Кваліфікаційна робота на тему «Інтерактивний застосунок для вивчення англійської мови на основі медіа-контенту із використанням OpenAI» на здобуття освітнього ступеня «бакалавр» зі спеціальності 122 «Комп'ютерні науки» освітньої програми «Комп'ютерні науки» написана обсягом в 69 сторінок і містить 33 ілюстрацій, 8 додатків та 27 використаних джерел.

Метою кваліфікаційної роботи є створення інтерактивного застосунку для вивчення іноземної мови на основі медіа-контенту із використанням штучного інтелекту.

Об'єктом дослідження є процеси отримання, зберігання, обробки, передачі та використання інформації.

Результатом роботи є багатофункціональний застосунок, що поєднує перегляд відеоконтенту, інтерактивну роботу з субтитрами, навчальні вправи та інструменти OpenAI для контекстного вивчення лексики, збереження її у словники й подальшого закріплення знань.

Розроблено інтерактивний застосунок для вивчення англійської мови на основі медіа-контенту із використанням OpenAI, який можуть використовувати школярі та студенти для покращення англійської мови з фокусом на розуміння живої мови, дорослі користувачі для вдосконалення мовних навичок у повсякденному, професійному або академічному середовищі, а також викладачі англійської мови, як додатковий дидактичний інструмент для інтерактивних занять і створення адаптивних завдань.

Ключові слова: ІНТЕРАКТИВНИЙ ЗАСТОСУНОК, ПЕРСОНАЛІЗОВАНЕ НАВЧАННЯ, ВИВЧЕННЯ АНГЛІЙСЬКОЇ МОВИ, МЕДІАКОНТЕНТ, СУБТИТРИ, ГЕНЕРАЦІЯ ТЕСТІВ, ШТУЧНИЙ ІНТЕЛЕКТ.

ANNOTATION

Qualification work on the topic «An interactive application for learning English based on media content using OpenAI» for obtaining a bachelor's degree in the specialty 122 «Computer Science» of the educational program «Computer Science» is written on 69 pages and contains 33 illustrations, 8 appendices, and 27 references.

The aim of this qualification work is to develop an interactive application for learning a foreign language based on media content with the use of artificial intelligence.

The object of research is the processes of acquisition, storage, processing, transmission, and use of information.

The result of the work is a multifunctional application that combines video content viewing, interactive subtitle interaction, learning exercises, and OpenAI-based tools for contextual vocabulary learning, saving words to personal dictionaries, and reinforcing acquired knowledge.

An interactive application for learning English based on media content using OpenAI was developed. It can be used by schoolchildren and students to improve their English with a focus on understanding natural language, by adult users to enhance language skills in everyday, professional, or academic contexts, and by English teachers as an additional didactic tool for interactive lessons and creating adaptive tasks.

Keywords: INTERACTIVE APPLICATION, PERSONALIZED LEARNING, ENGLISH LANGUAGE LEARNING, MEDIA CONTENT, SUBTITLES, TEST GENERATION, ARTIFICIAL INTELLIGENCE.

ЗМІСТ

Вступ.....	7
1 Аналіз предметної області і постановка задачі проектування.....	10
1.1 Опис інтерактивного вивчення іноземної мови на основі медіа-контенту....	10
1.2 Огляд і аналіз існуючих аналогів інтерактивного вивчення іноземної мови на основі медіа-контенту	15
1.3 Постановка задачі проектування.....	18
2 Алгоритмічне та інформаційне забезпечення інтерактивного застосунку для вивчення іноземної мови на основі медіа-контенту із використанням ШІ	21
2.1 Загальна структура проєктованого інтерактивного застосунку	21
2.2 Алгоритмічне забезпечення інтерактивного застосунку	24
2.3 Інформаційне забезпечення інтерактивного застосунку	29
3 Реалізація інтерактивного застосунку для вивчення іноземної мови на основі медіа-контенту із використанням OpenAI.....	33
3.1 Програмна реалізація інтерактивного застосунку з використанням OpenAI	33
3.2 Інтерфейс користувача інтерактивного застосунку	37
3.3 Тестування інтерактивного застосунку для вивчення іноземної мови.....	42
Висновки	46
Список використаних джерел	48
Додаток А Схема алгоритму роботи із субтитрами	51
Додаток Б Схема алгоритму роботи із бібліотекою медіа-контенту	52
Додаток В Схема алгоритму інтерактивної взаємодії із субтитрами.....	53
Додаток Г Схема алгоритму роботи із бібліотекою лексики	54
Додаток Д Схема алгоритму вивчення лексики	55
Додаток Е Атрибути і первинні ключі сутностей БД	56
Додаток Ж Програмний код основного функціоналу застосунку	59
Додаток И Копія опублікованих результатів	65

ВСТУП

У теперішніх реаліях, у контексті активного розвитку цифрових технологій та обміну інформацією, все більша частина населення стикається із новими вимогами до комунікації та взаємодії між людьми. За таких обставин знання іноземних мов перестає бути лише додатковим навиком – володіння ними забезпечує доступ до глобальних ресурсів, створюючи нові можливості для навчання, кар'єрного зросту, ефективної комунікації у різних сферах діяльності і доступу до величезної кількості інформації без необхідності перекладу.

Актуальність теми полягає у тому, що на тлі сучасних тенденцій, де швидкий ритм життя, кліпове мислення та величезний обсяг інформації стали невід'ємними складовими, особливо важливим є знаходження ефективних способів обробки, відбору та засвоєння цієї інформації. На сьогоднішній день, переважна частина людей наштовхується на переповнення даними й одночасно шукає способи поєднання навчання та відпочинку. Це зумовлює потребу в автоматизації процесів навчання, що сприятиме максимально ефективному засвоєнню нових навичок без необхідності трати надмірно великої кількості часу та енергії. Поєднання комфорту, зручності й саморозвитку дозволяє створювати нові можливості для особистісного зростання, підвищуючи якість життя та забезпечуючи баланс між відпочинком і навчанням.

Завдяки доступу до Інтернету, вивчення іноземних мов стало більш різноманітним та доступним. Особливу популярність здобули методи із використанням медіа-контенту. Цей підхід дозволяє інтегрувати навчання у повсякденне життя, роблячи процес більш природним і цікавим. При чому саме інтерактивне вивчення кардинально відрізняється від традиційних методів, передбачаючи не лише пасивне споживання інформації, а й активну взаємодію з нею [1].

Кіномедіа є одним із найпопулярніших видів даного інтерактивного вивчення мови, адже здатний поєднувати розважальний контент із навчальними елементами. Воно є прогресивним, доступним та актуальним напрямом, оскільки дає можливість зануритися у мовне середовище через реальні ситуації, що значно покращує розуміння і сприйняття. Великий вибір на стримінгових платформах робить цей процес гнучким та доступним для широкої аудиторії. Більш того, наявність

інтерактивних функцій у застосунках, такі як субтитри, транскрипти, вправи та підказки, допомагає активно взаємодіяти з контентом, сприяє кращому засвоєнню матеріалу та підвищує мотивацію, викликаючи емоційний відгук і зацікавленість.

Однак є й певні обмеження і недоліки. Наприклад, немало платформ мають слабо розвинену функціональність для інтерактивного навчання чи обмежену базу даних медіа, що змушує користувачів шукати додаткові інструменти чи переходити між різними програмами. Крім того, ефективне навчання потребує доповнення іншими методами, такими як граматика, збагачення словникового запасу тощо.

Мета даної роботи полягає в розробці інтерактивного застосунку для вивчення іноземної мови на основі медіа-контенту із використанням штучного інтелекту (ШІ). Для досягнення поставленої мети необхідно виконати наступні завдання:

- аналіз предметної області;
- огляд існуючих аналогів і постановка задачі проектування;
- опис структури проєктованого інтерактивного застосунку для вивчення іноземної мови на основі медіа-контенту;
- розробка алгоритмічного та інформаційного забезпечення інтерактивного застосунку;
- розробка інтерфейсу та функціоналу додатку, що включає можливість перегляду медіа з інтерактивними елементами для вивчення мови;
- інтеграція дієвих інструментів вивчення мови, таких як субтитри, повторення, переклад, визначення нових слів та виразів в контексті;
- тестування різних форм інтерактивності для покращення процесу навчання, включаючи відео, тести, та оцінки;
- тестування застосунку та визначення можливих обмежень і недоліків при використанні.

Об'єктом дослідження є процеси отримання, зберігання, обробки, передачі та використання інформації.

Предметом дослідження є інтерактивний застосунок для вивчення англійської мови із використанням OpenAI.

Практичне значення отриманих результатів: розроблений продукт може служити корисним і зручним доповненням для вивчення іноземних мов через кіномедіа-контент із використанням сучасних можливостей ІІІ, поєднуючи ефективність опанування мови із комфортом користування.

Результати дослідження опубліковано в матеріалах науково-практичної конференції «Інтелектуальні інформаційні технології в прикладних дослідженнях» (ІІТАР – 2025), м. Тернопіль, 27–29 травня 2025 р. (додаток ІІ).

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ І ПОСТАНОВКА ЗАДАЧІ ПРОЕКТУВАННЯ

1.1 Опис інтерактивного вивчення іноземної мови на основі медіа-контенту

Вивчення іноземної мови на основі медіа-контенту – це підхід до мовного навчання, що використовує різноманітні відео, аудіо та текстові матеріали як основний засіб освоєння мови [2]. Медіа-контент у цьому контексті може включати фільми, серіали, документальні програми, відео на різноманітні тематики основною ідеєю яких є занурення в реальне мовне середовище через автентичні матеріали, що використовуються носіями мови.

Людські очі це найчутливіший орган сприйняття інформації навколишнього світу. Спостереження за чимось візуально привабливим робить процес сприйняття та запам'ятовування даних більш ефективним, тому не дивно, що серед різних форматів медіа-контенту перегляд фільмів і серіалів є одним із найпоширеніших способів вивчення мови – вони пропонують багатий контекст, різноманіття мовних ситуацій та автентичні діалоги, що сумарно допомагає в рази краще сприймати живу мову [3]. На відміну від суто аудіо-матеріалів або текстового контенту, відео поєднує візуальну та аудіальну складові, що сприяє глибшому розумінню мови та її використання в реальних комунікативних ситуаціях.

Ще однією унікальною складовою кіномедіа є мистецтво оповідання історії, що захоплює нашу увагу для передачі важливої інформації із наступним впливом на емоційну, мотиваційну та когнітивну складові людини [4]. Особливо це актуально у реаліях сьогодення, де у все більшій частині населення розвиненим стає кліпове мислення – тип мислення, який характеризується швидким сприйняттям та обробкою інформації і постійним переключенням уваги з одного об'єкта на інший, що виник через пришвидшення ритму життя, збільшення обсягів інформації і відповідних збільшеннях потреб у її поглинанні та раціональному відсіюванні [3, 5]. У контексті кіномедіа оповідання історії може позитивно впливати на:

- поглиблення емоційного зв'язку з персонажами чи створення певного емоційного відгуку загалом, а все, що викликає у людини емоції, в результаті набагато краще відбивається у пам'яті [6];

— підвищення концентрації уваги завдяки певній динаміці — мотивація через зацікавленість до сюжету здатна викликати зосередженість у слуханні й аналізуванні мови навіть у моменти, коли слова або фрази важко зрозуміти. Отже, сюжет стає не тільки носієм емоційної складової, але й інструментом, що допомагає мотивувати глядача [7];

— постійне повторення та закріплення лексики через частіше зустрічання певних фраз або слів у різних епізодах чи ситуаціях, що є кращим підходом, аніж просте механічне запам'ятовування.

Створення ефективних освітніх умов забезпечує природну практику занурення у мовне середовище, що дає можливість особі, яка її опановує, взаємодіяти із лінгвістичним матеріалом слухаючи, читаючи та працюючи із справжніми життєвими обставинами, що власне і робить вивчення слів, виразів, фразеологізмів, цілих мовленнєвих зворотів набагато легшим [8].

Окрім того, даний стиль навчання допомагає підтримувати мотивацію — цікаві матеріали з орієнтацією на індивідуальні потреби й смаки заохочують продовжувати рухатись у прогресі, перетворюючи сам освітній процес у радше захопливий вид діяльності без зовнішнього тиску, аніж у марафон [9].

Цілі підходу включають значне покращення навичок аудіювання та розуміння розмовної мови, адже користувачі здобувають досвід слухання носіїв у реальних комунікаційних ситуаціях, таких як бесіди, інтерв'ю тощо. Таке навчання сприяє активному розширенню словникового запасу завдяки контекстуальному сприйняттю нових слів і виразів, а сюжет та емоційне тло допомагають краще зрозуміти нюанси їх використання у різних соціальних ситуаціях.

Правила використання граматичних конструкцій сприймаються безпосередньо в процесі прослуховування реального мовлення. Це дозволяє краще розуміти застосування граматичних форм на практиці [10]. Також це дає більш глибоке ознайомлення з культурним контекстом, акцентами, особливостями вимови, специфічною лексикою, що є важливим аспектом даного підходу і робить процес навчання ще ефективнішим.

В основному, більшість користувачів застосовують пасивний метод навчання — переглядають контент без активного втручання, а основний фокус

відводиться на аудіювання та розуміння контексту. Пасивні методи хоча й можуть бути корисними для занурення у мовне середовище, але часто є недостатньо ефективними через низку причин.

По-перше, перегляд відео без активної участі не сприяє закріпленню знань, оскільки мозок обробляє інформацію поверхнево, не аналізуючи граматичні конструкції та нові слова. Без активного повторення та використання вивченого матеріалу засвоєння мови значно уповільнюється.

По-друге, субтитри, хоча й можуть допомогти в розумінні контенту, часто перетворюються на «опору», через що користувач більше концентрується на читанні, а не на розпізнаванні слів на слух. Це може створювати хибне відчуття розуміння мови, тоді як насправді при розмові людина все ще відчуватиме труднощі.

Крім того, пасивний перегляд не розвиває продуктивні мовні навички, такі як говоріння та письмо, які є ключовими для повноцінного володіння мовою. Без практичного застосування знань у вигляді інтерактивних завдань або вправ користувач лише пасивно сприймає інформацію, не формуючи стійких мовних вмінь.

Тому доцільним є створення ресурсу, який б поєднував перегляд кіномедіа-контенту з активними методами навчання: інтерактивними вправами, тестами на закріплення матеріалу, можливістю відстеження та аналізу власного прогресу, а також простого й інтуїтивно зрозумілого у використанні інтерфейсу із широкими можливостями користувацьких налаштувань. Такий підхід дозволить не лише отримувати задоволення від перегляду, а й ефективно інтегрувати вивчене у власне використання та застосовувати отримані знання на практиці, що значно прискорить процес опанування мови із можливістю комфортної адаптації навчання.

Основними перевагами стануть:

- активне засвоєння матеріалу – користувач не просто переглядає контент, а взаємодіє з ним через інтерактивний перегляд, тести, що сприяє більш глибокому опрацюванню, а значить і кращому запам'ятовуванню. Також це підвищує залученість у процес;
- розвиток усіх мовних навичок – система сприятиме комплексному опануванню мови, розвиваючи не лише покращене її розуміння на слух, а й інші ключові

аспекти, такі як письмо і читання, завдяки чому можна створити збалансоване навчальне середовище, покращуючи якість засвоєння матеріалу;

- контроль і відстеження прогресу – можливість переглядати власні досягнення у динаміці, відстежування та аналіз результатів виконаних вправ через графіки допоможуть підтримувати мотивацію та дозволять користувачу легше коригувати стратегію навчання відповідно до особистих потреб;

- гнучкість навчання – користувач отримує можливість самостійно обирати контент, який відповідає його інтересам, рівню знань та навчальним цілям. Це дозволяє уникнути фіксованих дедлайнів, жорстких рамок або неактуальних матеріалів і навчатися у зручному темпі;

- краще засвоєння вимови та інтонації – робота з автентичним медіа-контентом дозволяє чути природну мову такою, якою вона використовується в реальному житті, що служить не лише правильному сприйняттю вимови, а й освоєнню інтонаційних особливостей, ритму, різноманітних акцентів завдяки постійному контакту з живими діалогами та розмовними фразами;

- підвищення ефективності навчання – поєднання розважального контенту з навчальними елементами сприятиме збереженню інтересу і створенню сприятливого середовища для засвоєння нової інформації. Завдяки такій комбінації відпочинку та самоосвіти, навчальний процес стає більш невимушеним і легким, що зменшує когнітивне навантаження та значно пришвидшує ріст рівня володіння мовою природним шляхом.

Проте на цьому можливості такого підходу не зупиняються, оскільки сучасні тенденції розвитку технологій надали нам ще ширші горизонти для вдосконалення користувацького досвіду через інтеграцію штучного інтелекту (ШІ). Його застосування здатне автоматизувати створення тестів для самоконтролю задля закріплення матеріалу по прослуховуванню після перегляду певного медіа-контенту.

Отже, пасивне вивчення, безумовно, приносить чималу користь, поглиблюючи розуміння мови і звикання до її звучання, однак інтерактивні методи навчання, що включають закріплення матеріалу через вправи та тести, у поєднанні з додатковими можливостями ШІ, значно підвищують ефективність навчального процесу, забезпечуючи індивідуальний підхід та постійну адаптацію до потреб користувача.

Функціональна структура бізнес-процесу інтерактивного застосунку для вивчення іноземної мови на основі медіа-контенту з використанням ШІ включає такі елементи, як: стандартизація, раціональне використання ресурсів, ефективна комунікація, гнучкість і адаптивність. Ієрархічне представлення цієї структури наведено на рисунку 1.1.



Рисунок 1.1 – Функціональна структура інтерактивного застосунку для вивчення іноземної мови на основі медіа-контенту із використанням ШІ

Запропонована функціональна структура відображає комплексний підхід до вивчення іноземної мови, поєднуючи різні методи і засоби навчання: перегляд медіа-контенту, інтерактивну роботу з субтитрами, збереження та повторення лексичного матеріалу, аналітику динаміку прогресу, інтеграцію з браузером через розширення та персоналізоване тестування завдяки використанню засобів ШІ із врахування контексту та мовного наповнення.

1.2 Огляд і аналіз існуючих аналогів інтерактивного вивчення іноземної мови на основі медіа-контенту

У сучасному цифровому середовищі інтерактивні застосунки для вивчення іноземних мов стали важливим інструментом для мільйонів користувачів, поєднуючи мультимедійні можливості та адаптивні методики навчання. Для виявлення ефективних підходів було проведено аналіз існуючих аналогів, які у тій чи іншій мірі успішно реалізують подібний функціонал для ефективного опанування мов, а саме застосунок для вивчення англійської та корейської мов «Cake» і веб-ресурс для вивчення англійської на основі медіа-контенту «Listen in English».

«Cake» – це популярний інструмент для вивчення англійської та корейської мов. Він пропонує різноманітний функціонал, який допомагає користувачам покращити свої мовні навички [11]:

- користувачам щодня пропонуються нові мовні вирази та навчальні матеріали, що включають відео за участю знаменитостей, акторів, музикантів й популярних відеоблогерів;
- надається можливість активно тренувати мовлення та навички аудіювання шляхом проходження інтерактивних вікторин і використання інструментів автоматичної перевірки вимови, реалізованих на основі технологій ШІ;
- моделі покрокового навчання пропонують структуровані курси, розроблені експертами у галузі лінгвістики, що дозволяє вивчати матеріал поступово у власному темпі;
- можливість вивчати справжні мовні конструкції, які використовуються носіями англійської чи корейської у повсякденних ситуаціях – зокрема у влогах, художніх фільмах, серіалах і телешоу;
- мотивація користувачів підтримується засобами інтерактивних елементів у вигляді рейтингів, досягнень і змагальних механік, що дозволяє відстежувати власний прогрес і залишатися зацікавленими та цілеспрямованими упродовж усього процесу навчання.

Система представлена у форматі веб-ресурсу та у вигляді мобільного застосунку. Має адаптивний дизайн й інтуїтивно зрозумілий інтерфейс (рисунки 1.2).

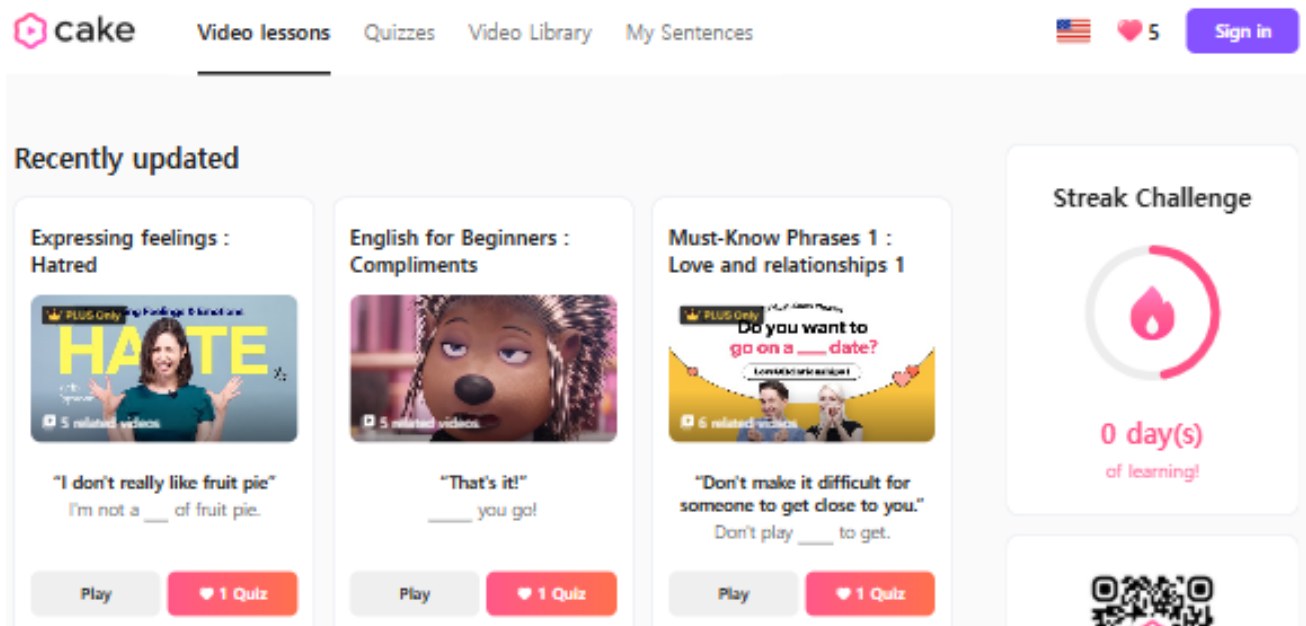


Рисунок 1.2 – Головна сторінка застосунку «Cake»

Отож, даний застосунок має низку переваг та зручностей для навчання, зокрема він поєднує різноманіття щоденного медіа-контенту, інтерактивних завдань, різних напрямків відточення навичок (читання, аудіювання, письмо, вимова), а також застосовує можливості ШІ для покращення користувацького досвіду. Проте, не зважаючи на це, існують подальші можливості для вдосконалення:

- інтерактивна взаємодія із субтитрами;
- інтеграція із певним стримінговими сервісами без порушення авторських прав для повного доступу до медіа-контенту (зокрема напрямку кіно);
- більш адаптивна інтервальна система повторення та гнучкіша система створення і керування персональними словниками (категоріями).

«Listen in English» – це вебсайт орієнтований на ефективне вивчення англійської мови. Даний популярний вебресурс був розроблений спеціально для тих, хто прагне покращити свої навички сприйняття мови на слух. Платформа пропонує широкий спектр інтерактивних вправ, які ґрунтуються на оригінальних матеріалах із художніх фільмів, телесеріалів, новинних відео-сюжетів, інтерв'ю та подкастів, що допомагає максимально наблизитись до реального мовного середовища.

Ресурс функціонує виключно у вебверсії, має інтуїтивно зрозумілий інтерфейс із простою навігацією та хорошою структурованістю матеріалів (рисунок 1.3).

1. Monica and Phoebe are angry because Ross kissed Rachel.
☐ True
☐ False
2. Phoebe tells Monica to make a phone call.
☐ True
☐ False
3. The girls want to drink wine while they are listening to Rachel tell the story.
☐ True
☐ False

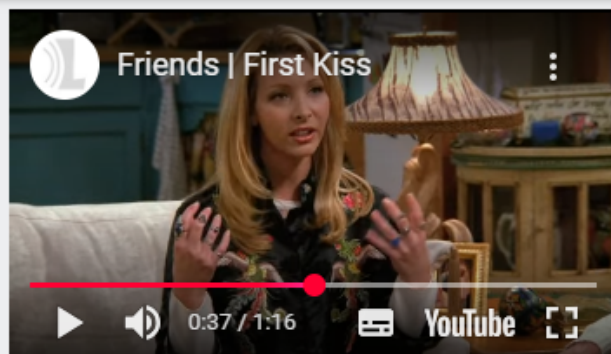


Рисунок 1.3 – Сторінка «Навчання» вебресурсу «Listen in English»

Одною із головних переваг є можливість вибору рівня складності. Всі матеріали структуровані відповідно до загальноєвропейської системи рівнів володіння мовою (CEFR), що дозволяє кожному користувачеві навчатися у власному темпі. Сайт пропонує завдання для початківців, середнього рівня та просунутих учнів, що робить його корисним як для самостійного навчання, так і для онлайн-уроків [12].

Крім цього, платформа дає змогу тренувати сприйняття різних акцентів та мовних ситуацій, пропонуючи завдання, які допоможуть звикнути до звучання англійської у реальному житті – від повсякденних розмов до виступів носіїв мови.

Проте даний вебресурс все ж має певні обмеження, компенсація яких могла би якісно вплинути на користувацький досвід і процес навчання:

- відсутність повноцінного медіа-контенту: одним із недоліків, з яким неочікувано може зіткнутися користувач є те, що ресурс пропонує лише короткі фрагменти відео та аудіо (на кілька хвилин), що обмежує можливість повного занурення в мовне середовище. Загалом це може стати доволі незручним для користувачів, що планують заглибитися у фільм чи серіал повною мірою і вивчати мову дивлячись матеріал цілком, а не фрагментами чи вирізками;

- доволі невисокий рівень інтерактивності: платформа пропонує більш пасивне споживання контенту – користувачі можуть слухати аудіо чи дивитися фрагменти відео, але не можуть взаємодіяти з ними, зокрема для внесення слів чи фраз до персональних категорій словників прямо в момент взаємодії із контентом.

Отже, доцільно розробити продукт, що матиме на меті подолати основні недоліки існуючих рішень: нестачу інтерактивності (зокрема при самому перегляді), обмежений медіа-контент (повноцінний доступ до розширеного контенту замість коротких фрагментів відео) та відсутність персоналізації (можливість створення власних тематичних підбірок).

1.3 Постановка задачі проектування

З огляду на проведені у попередніх підрозділах огляд та аналіз предметної області й існуючих аналогів, проєктований застосунок повинен забезпечувати комплексну автоматизацію процесу вивчення іноземної мови, тобто впровадження технологій та функціональних можливостей, які спрощують та оптимізують навчання, та бути надійним, масштабованим і зручним у використанні, пропонуючи гнучкі навчальні плани для ефективного засвоєння матеріалу.

Мета даної роботи полягає в розробці інтерактивного застосунку для вивчення іноземної мови на основі медіа-контенту із використанням штучного інтелекту (ШІ). Для досягнення поставленої мети необхідно виконати наступні завдання:

- аналіз предметної області;
- огляд існуючих аналогів і постановка задачі проектування;
- опис структури проєктованого інтерактивного застосунку для вивчення іноземної мови на основі медіа-контенту;
- розробка алгоритмічного та інформаційного забезпечення інтерактивного застосунку;
- розробка інтерфейсу та функціоналу додатку, що включає можливість перегляду медіа з інтерактивними елементами для вивчення мови;
- інтеграція дієвих інструментів вивчення мови, таких як субтитри, повторення, переклад, визначення нових слів та виразів в контексті;
- тестування різних форм інтерактивності для покращення процесу навчання, включаючи відео, тести, та оцінки;

- тестування застосунку та визначення можливих обмежень і недоліків при використанні.

Нижче сформовано задачі до застосунку:

- створення механізму інсталяції, керування та деінсталяції розширення у браузері;
- розробити інтеграції з популярними стримінговими платформами та створити сторінку з каталогом підтримуваних сервісів;
- розробити функціонал для завантаження та відображення субтитрів. створити редактор користувацьких субтитрів, реалізувати функції налаштування параметрів відображення субтитрів;
- створити сховище для зберігання переглянутого матеріалу. розробити інтерфейс відображення медіа-бібліотеки та реалізувати функцію запам'ятовування часової мітки зупинки перегляду, додати можливість сортування та фільтрації контенту в бібліотеці;
- розробити функціонал для створення, редагування та видалення категорій лексики;
- створити інтерфейс для керування словами в кожній категорії. реалізувати функції додавання, копіювання та видалення лексичних одиниць;
- інтегрувати систему генерації аудіо-тестів з використанням ШІ і розробити алгоритм періодичного повторення лексики;
- розробити систему збору статистичних даних про процес навчання; створити панель аналітики з різними видами графіків та діаграм; реалізувати функцію генерації звітів про прогрес;
- розробити механізм створення резервних копій даних користувача; створити інтерфейс для керування резервними копіями і реалізувати функцію відновлення даних з резервної копії.

Отож, процеси, що оптимізуються перелічені нижче:

- управління медіаконтентом, що передбачає можливість зберігати, організовувати та категоризувати дані, відстежувати прогрес перегляду та здійснювати фільтрацію за різними параметрами;
- опрацювання субтитрів, їх підключення, синхронізація та переклад;

- формування та поповнення лексичного запасу із можливістю миттєвого зберігання нових слів та виразів і їх наступною організацією у персональні словники для подальшого засвоєння через інтервальну систему повторень;
- персоналізоване навчання із використанням ШІ для генерації тестів на основі переглянутого контенту;
- моніторинг та аналіз прогресу навчання за допомогою графіків і діаграм;
- інтеграція з веб-браузером;
- резервне копіювання даних та управління локальним сховищем.

Щодо інформаційного забезпечення, застосунок повинен містити базу даних (БД), що включає інформацію про медіа-контент, доступні джерела для перегляду, слова (вирази, фрази), словники (категорії матеріалів), завдання (тести, оцінки), базові дані користувача (налаштування, сховище даних).

Запропонований застосунок здатний суттєво підвищити ефективність вивчення іноземної мови порівняно з традиційними методами завдяки таким рішенням:

- контекстуальне засвоєння лексики на відміну від ізольованого вивчення слів за словниками;
- природний мовний контекст;
- інтеграція різних каналів сприйняття із одночасним задіянням аудіальної та візуальної складових із підтримкою текстових субтитрів, що створює мультисенсорне середовище;
- персоналізація навчального процесу із можливістю адаптації інтервального повторення слів і виразів відповідно до індивідуальних потреб;
- автоматизація рутинних процесів, що скоротить технічні затрати часу, дозволяючи зосередитись безпосередньо на навчанні;
- інтерактивність взаємодії з контентом, що трансформує пасивний перегляд у активний навчальний процес;
- аналітика прогресу і його візуалізація через графіки та діаграми;
- ШІ-адаптація навчальних матеріалів для генерації персоналізованих тестів на основі вже переглянутого контенту, що забезпечить релевантність завдань.

2 АЛГОРИТМІЧНЕ ТА ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ ІНТЕРАКТИВНОГО ЗАСТОСУНКУ ДЛЯ ВИВЧЕННЯ ІНОЗЕМНОЇ МОВИ НА ОСНОВІ МЕДІА-КОНТЕНТУ ІЗ ВИКОРИСТАННЯМ ІІІ

2.1 Загальна структура проєктованого інтерактивного застосунку

Розроблюване рішення міститиме чітко виражені компоненти з власними атрибутами та методами, а саме:

- працюватиме з різними типами об'єктів, такими як медіа-файли, субтитри, словники, тести тощо;
- існуватимуть складні зв'язки між сутностями, наприклад, «відео має субтитри», «користувач додає слова у словник, проходить тести»;
- реалізація інкапсуляції даних (дані не змінюються безпосередньо ззовні, а доступ до них здійснюється через визначені методи) та їх логічної структури.

Тобто передбачається створення явно визначених сутностей (об'єктів), які взаємодіятимуть між собою, а усі ключові компоненти зручно моделюються у вигляді класів, що дозволить легко розширювати та підтримувати систему. З огляду на це, основним обрано саме об'єктно-орієнтований підхід (ООП), оскільки:

- структура даних є ключовим аспектом функціонування системи;
- взаємодія між сутностями має багаторівневий та динамічний характер;
- система передбачає можливість масштабування та розширення (додавання нових типів контенту, інтеграція з іншими платформами).

Це забезпечить структурну цілісність, розширюваність і підтримку складної ієрархії функціональних можливостей системи. Виходячи із цього, для повноти моделювання будуть використовуватися різні види UML-діаграм, що дозволить наочно представити як статичну структуру системи, так і динамічні процеси взаємодії її компонентів.

Для наочного представлення динаміки взаємодії між основними компонентами системи була створена UML-діаграма послідовностей (рисунок 2.1). Вона дозволяє ознайомити із основним порядком виклику методів та обміну даними між об'єктами на різних етапах виконання основних функцій системи [13].

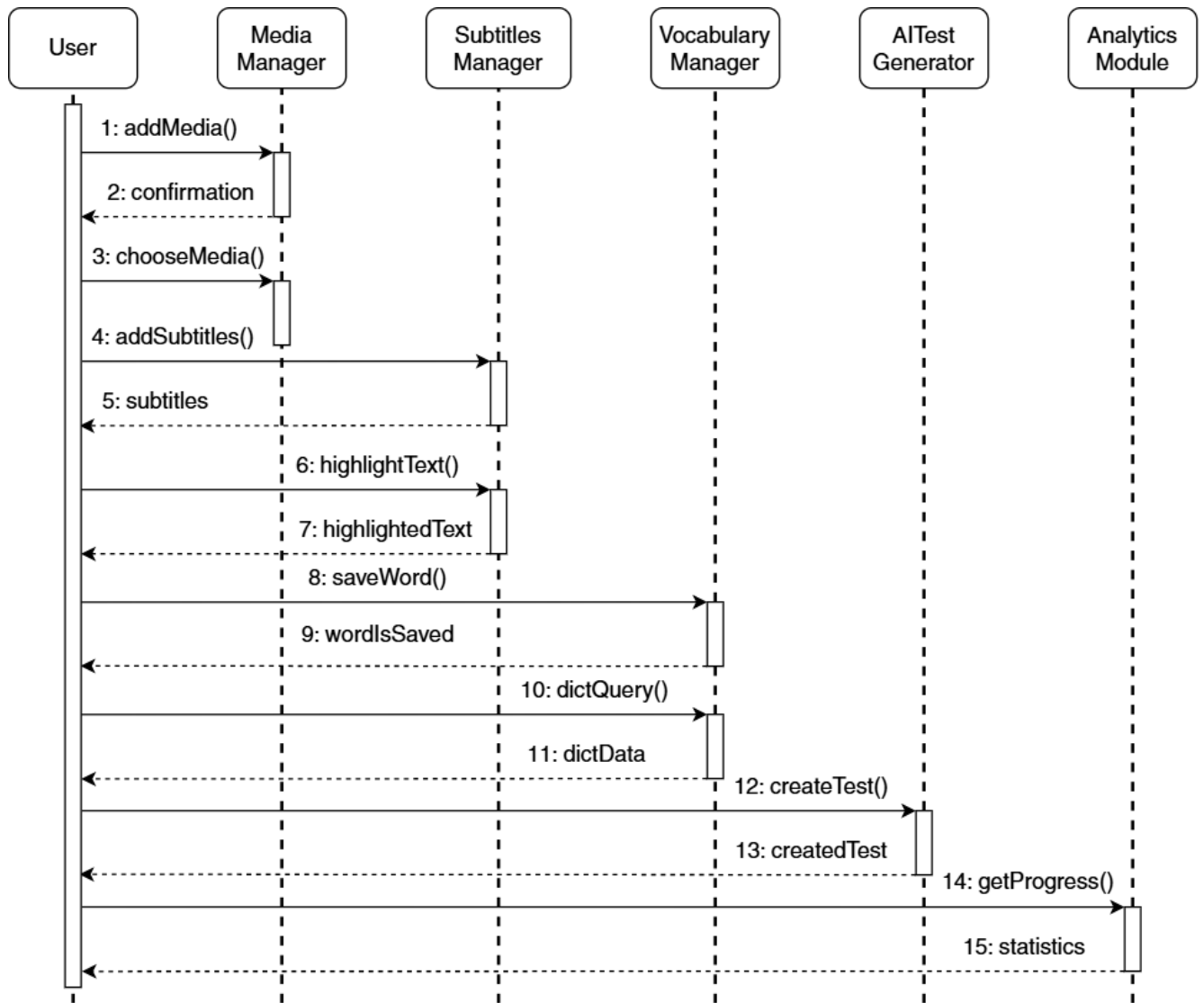


Рисунок 2.1 – UML діаграма послідовностей інтерактивного застосунку для вивчення англійської мови на основі медіа-контенту із використанням ШІ

Діаграма послідовності відображає основні взаємодії між користувачем та головними компонентами застосунку під час виконання типових дій, таких як додавання медіа-контенту, робота з субтитрами, збереження лексики та генерація тестів. Кожна з цих взаємодій продемонстрована на діаграмі як послідовність кроків, що включають обробку запитів користувача та підтвердження виконання дій відповідними модулями.

Для наочного представлення структури розроблюваного застосунку та взаємозв'язків між його основними компонентами була створена UML-діаграма класів (рисунок 2.2) – інструмент для опису статичної структури застосунку та взаємодії його складових на рівні класів об'єктно-орієнтованого підходу, який відображає їх атрибути, методи, а також зв'язки між ними [14].

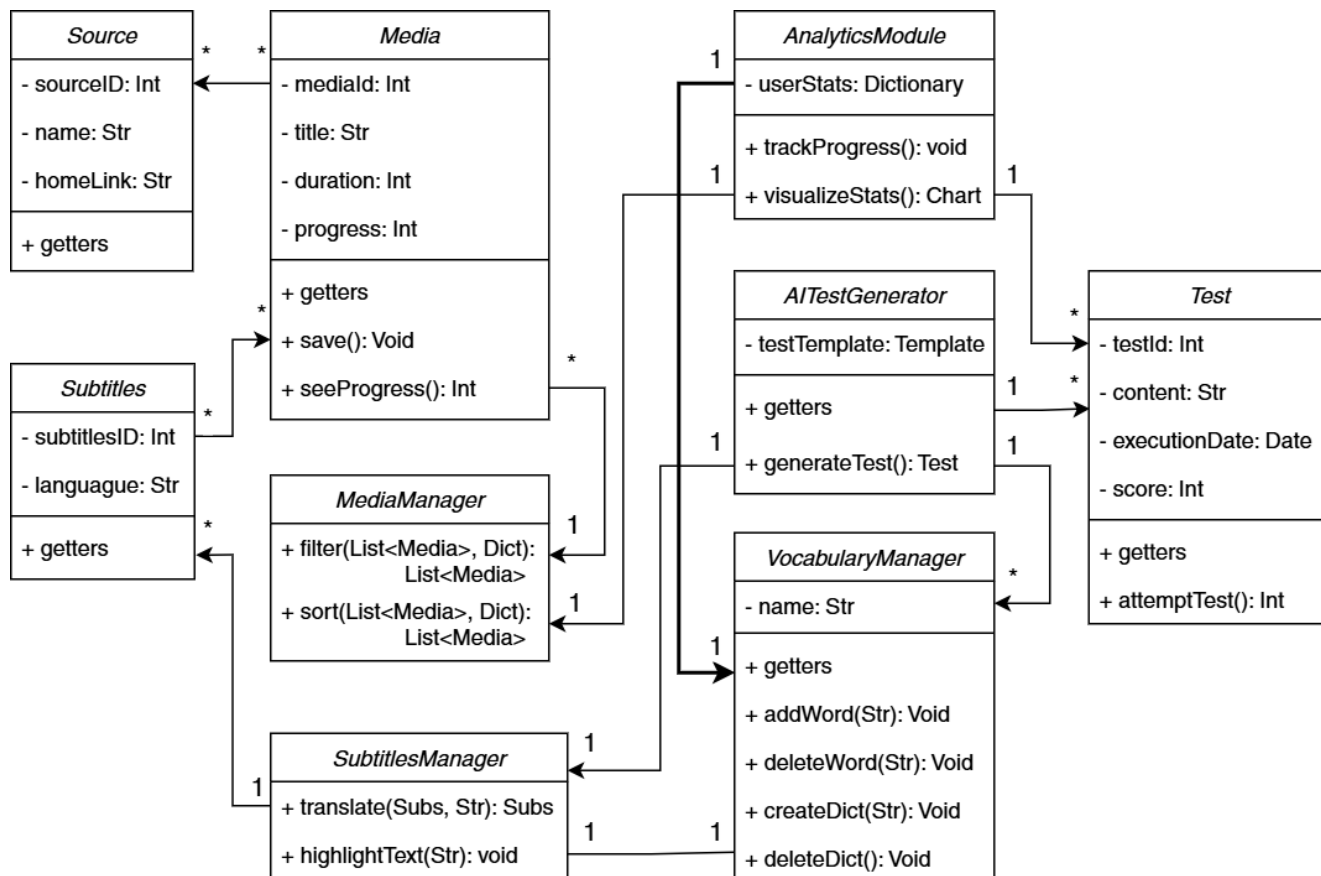


Рисунок 2.2 – UML діаграма класів інтерактивного застосунку для вивчення англійської мови на основі медіа-контенту із використанням OpenAI

Подана UML-діаграма (див. рисунок 2.2) визначає задачі класів:

- «Media» керує медіа-контентом, взаємодіючи з «Source» (джерелом) та «MediaManager» (сортування та фільтрація);
- «Subtitles Manager» обробляє субтитри, інтегруючись із «Vocabulary Manager» для збереження нових слів;
- «AI Test Generator» генерує тести, використовуючи дані субтитрів і словникового запасу;
- «Analytics Module» аналізує прогрес, взаємодіючи з «Media Content Manager» та «Vocabulary Manager».

Використання поданих діаграм на етапі проєктування допоможе забезпечити структурну цілісність, спростить подальші процеси реалізації та тестування, а також надасть можливість легко вносити зміни у разі необхідності. Таким чином, ці інструменти є важливою частиною розробки, що сприяє створенню надійного, масштабованого та ефективного програмного продукту.

2.2 Алгоритмічне забезпечення інтерактивного застосунку

На рисунку 2.3 зображено схему алгоритму роботи застосунку для інтерактивного вивчення мови на основі медіа-контенту із використанням ШІ [26].

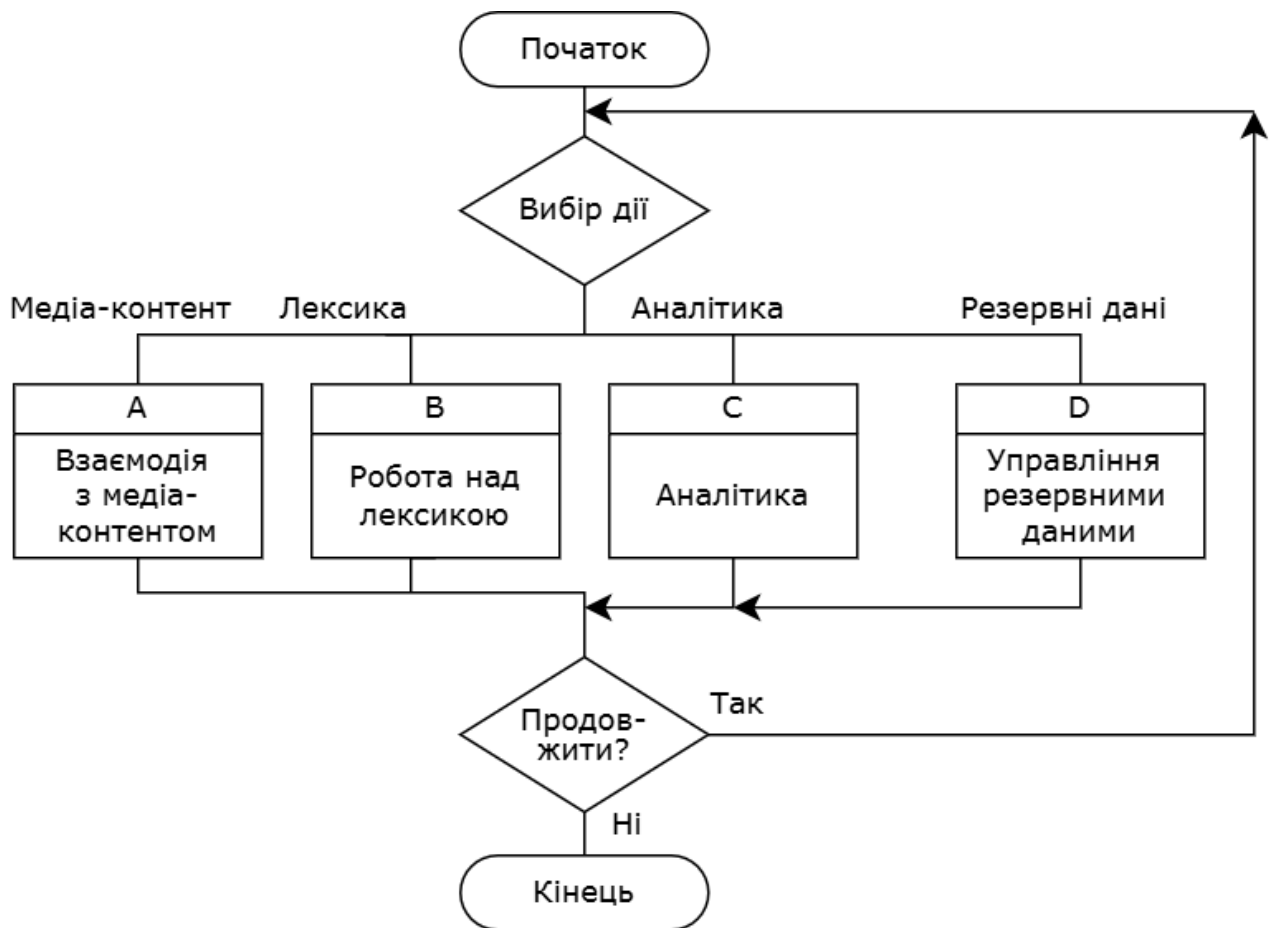


Рисунок 2.3 – Схема алгоритму роботи
інтерактивного застосунку для вивчення англійської мови

Із схеми алгоритму (див. рисунок 2.3) можна побачити наступну послідовність дій роботи застосунку:

1. вибір дії, яку потрібно виконати;
2. виконання обраної дії;
3. перевірка продовження користування; якщо «так», перехід на «1»;
4. завершення алгоритму.

Тепер розглянемо детальніше види діяльності, що доступні користувачу: взаємодія з медіа-контентом, робота над лексикою, аналітика та управління резервними даними.

На рисунку 2.4 зображено схему алгоритму взаємодії з медіа-контентом.

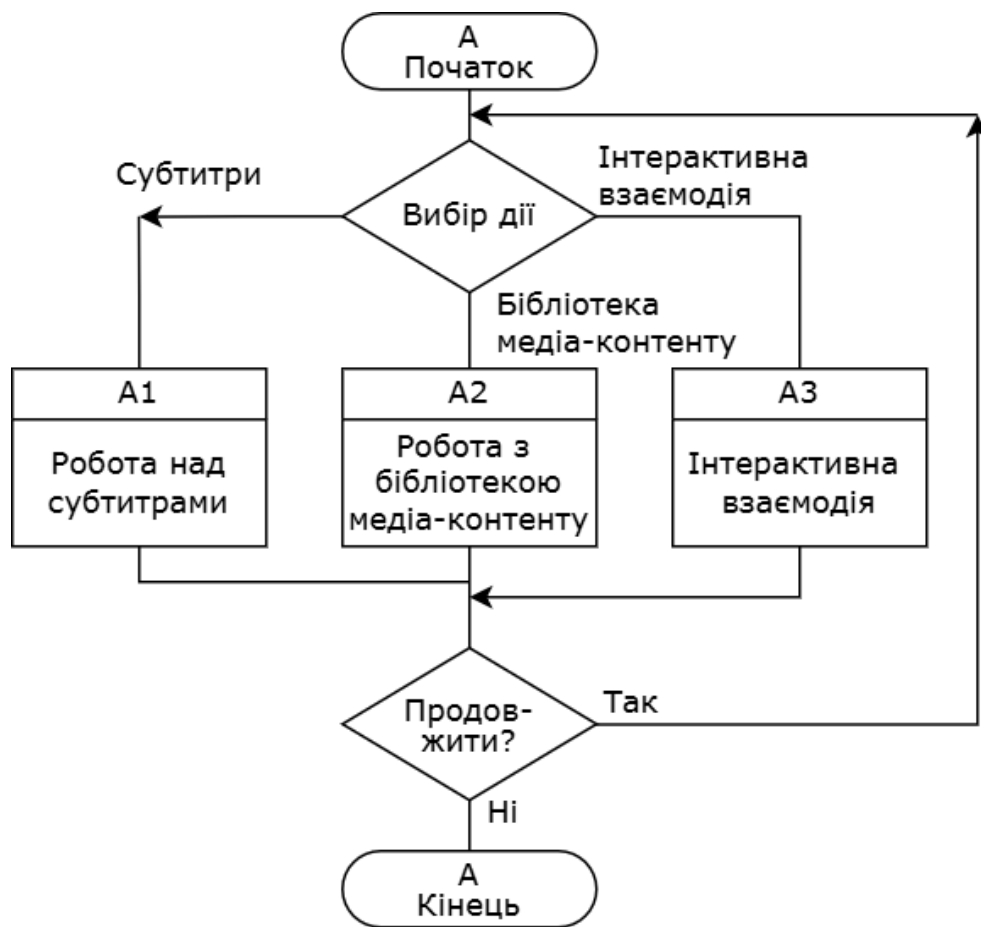


Рисунок 2.4 – Схема алгоритму модуля роботи з медіа-контентом

При роботі із медіа-контентом користувачу доступні такі можливості:

- робота над субтитрами (схема алгоритму представлена у додатку А);
- робота із збереженим контентом у бібліотеці медіа (схема алгоритму представлена у додатку Б);
- інтерактивна взаємодія із субтитрами при перегляді медіа (схема алгоритму представлена у додатку В).

Нижче розглянуто детальніше кожен із представлених вище пунктів.

Алгоритм роботи модуля із бібліотекою медіа-контенту покликаний вирішувати такі потреби користувача:

- додавання і видалення контенту із БД;
- фільтрація та сортування за різними критеріями для зручності навігації;
- перевірка валідності посилання на контент та обробка помилок;
- безперервний процес роботи з медіаконтентом до завершення сесії.

Алгоритм системного модуля роботи із субтитрами вирішує задачі:

- створення нових субтитрів на основі наявних текстових доріжок із стрімінгових сервісів із можливістю автоматичної генерації через Google Translate або MyMemory Translate;
- гнучке налаштування відображення субтитрів при перегляді (наприклад, одночасний показ оригіналу та перекладу);
- збереження субтитрів у БД для повторно використання у майбутньому;
- циклічний процес обробки переходу між діями.

Алгоритм модуля роботи інтерактивної взаємодії із субтитрами при перегляді медіа-контенту надає можливість виділяти та миттєво зберігати обрану лексику у користувацьку категорію.

Модуль роботи над лексикою забезпечує можливість ефективного засвоєння нових слів та виразів й організацію лексичного контенту для вивчення. Схема алгоритму представлена на рисунку 2.5.

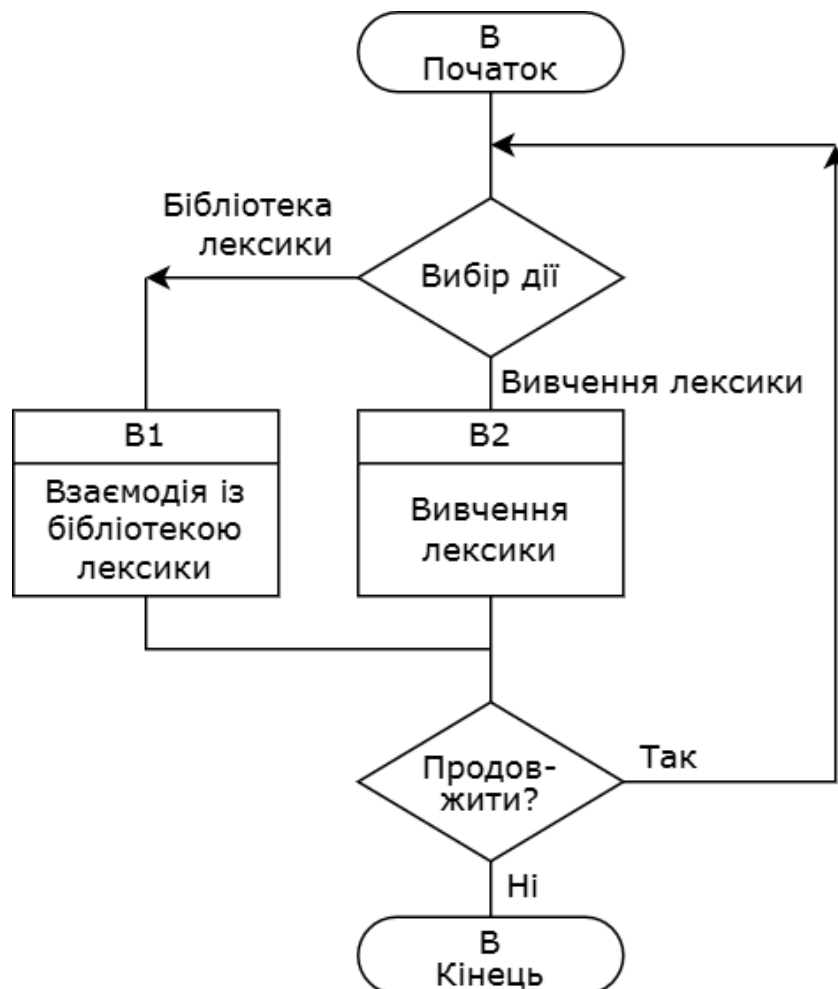


Рисунок 2.5 – Схема алгоритму модуля роботи із лексикою

При роботі над лексикою користувачу доступні такі варіанти використання:

- взаємодія із бібліотекою лексики – організація слів і виразів для навчання, зокрема через персональні категорії (схема алгоритму представлена у додатку Г);
- вивчення лексики – можливість поповнення словникового запасу через вбудовані методи застосунку (схема алгоритму представлена у додатку Д).

Модуль взаємодії із бібліотекою лексики призначений для повноцінного керування персональними категоріями слів (створення, видалення, редагування) та ефективної організації мовного матеріалу (добавлення, видалення, редагування, переніс чи копіювання між різними категоріями). Це здатне забезпечити гнучку роботу із лексичним контентом відповідно до індивідуальних потреб.

Модуль вивчення лексики орієнтований на всебічне покращення мовних здібностей, акцентуючи увагу на розвиток знань із аудіювання (реалізується шляхом проходження адаптивних аудіо-тестів, згенерованих ШІ на основі вже переглянутого медіа-контенту) та вимови і письма (здійснюється через перегляд інтерактивних карток на основі системи інтервальних повторень).

Модуль аналітики здійснює комплексне відстеження та аналіз прогресу навчання користувача. Схема алгоритму представлена на рисунку 2.6.

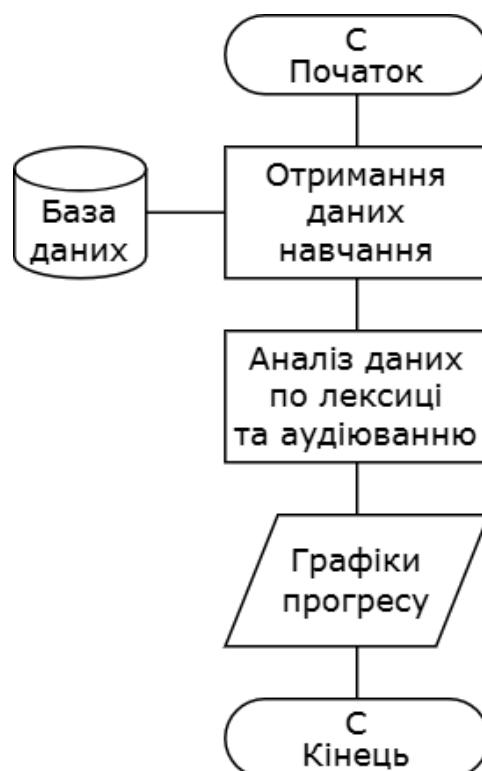


Рисунок 2.6 – Схема алгоритму модуля аналітики

Цей модуль покликаний полегшити процеси оптимізації та адаптації навчальної діяльності під індивідуальні потреби:

- аналіз наявних звітів включає етап збору та обробки даних, які зберігаються у БД. Дані класифікуються за типами навчальних завдань. Далі обчислюються базові метрики ефективності;
- побудова графіків здійснюється по лексичному прогресу, кількості помилок в аудіюванні, статусу навчання по словниках;
- результати надаються у легкому для розуміння форматі.

Модуль управління резервними даними відповідає за створення резервних копій і можливість відновлення даних історії прогресу. Схема алгоритму представлена на рисунку 2.7.



Рисунок 2.7 – Схема алгоритму модуля управління резервними даними

Описане алгоритмічне забезпечення дає чітке розуміння логіки роботи застосунку для проєктування коректної структури його БД.

2.3 Інформаційне забезпечення інтерактивного застосунку

Початково було здійснено етап концептуального проєктування предметної області (ПО), метою якого є визначення ключових сутностей (об'єктів реального світу) та їх взаємозв'язків без прив'язки до конкретної системи управління базами даних (СУБД). Проєктована база даних (БД) слугуватиме централізованим сховищем для збереження та управління інформацією платформ перегляду, медіа-контенту, лексичного матеріалу, завдань для тестування та аналітики навчального процесу користувача.

Список основних сутностей: джерело, медіа-контент, аудіо доріжка, жанр, країна, субтитри, мова, одиниця лексики, приклади використання лексики, категорія (словник), аудіо-тест. Нижче наведений опис зв'язків між сутностями:

- користувач може взаємодіяти із медіа-контентом на усіх доступних в застосунку джерелах;
- користувач може взаємодіяти із всіма субтитрами;
- користувач може генерувати та зберігати довільні варіації субтитрів;
- користувач може зберігати безліч медіа;
- медіа-контент може бути на різних ресурсах;
- медіа-контент може мати багато аудіо доріжок, субтитрів, жанрів та країн;
- користувач може створювати безліч персональних категорій слів;
- користувач може зберігати безліч одиниць лексики у категорії слів;
- кожна одиниця лексики може повторюватися у різних категоріях;
- кожна лексична одиниця може мати кілька прикладів її використання, і водночас один приклад може демонструвати застосування кількох лексичних одиниць;
- користувач може генерувати безліч аудіо-тестів і проходити їх довільну кількість разів.

Для створення на основі концептуальної структури логічної моделі даних необхідно усунути зв'язки «багато до багатьох» (досягається шляхом введення проміжних таблиць) та провести нормалізацію даних (для уникнення їх дублювання та аномалій при вставці, оновленні і видаленні). Розглянемо рисунок 2.8.



Рисунок 2.8 – Сутності «Медіа-контент», «Джерело», «Жанр» та «Країна» із наявними зв’язками «багато до багатьох»

Необхідно подолати наявні зв’язки «багато до багатьох» між «Медіа-контентом» та іншими трьома сутностями (див. рисунок 2.8). Для цього введемо проміжні таблиці «Джерела медіа-контенту», «Жанри медіа-контенту» та «Країни медіа-контенту». Отримано наступну структуру (рисунок 2.9).

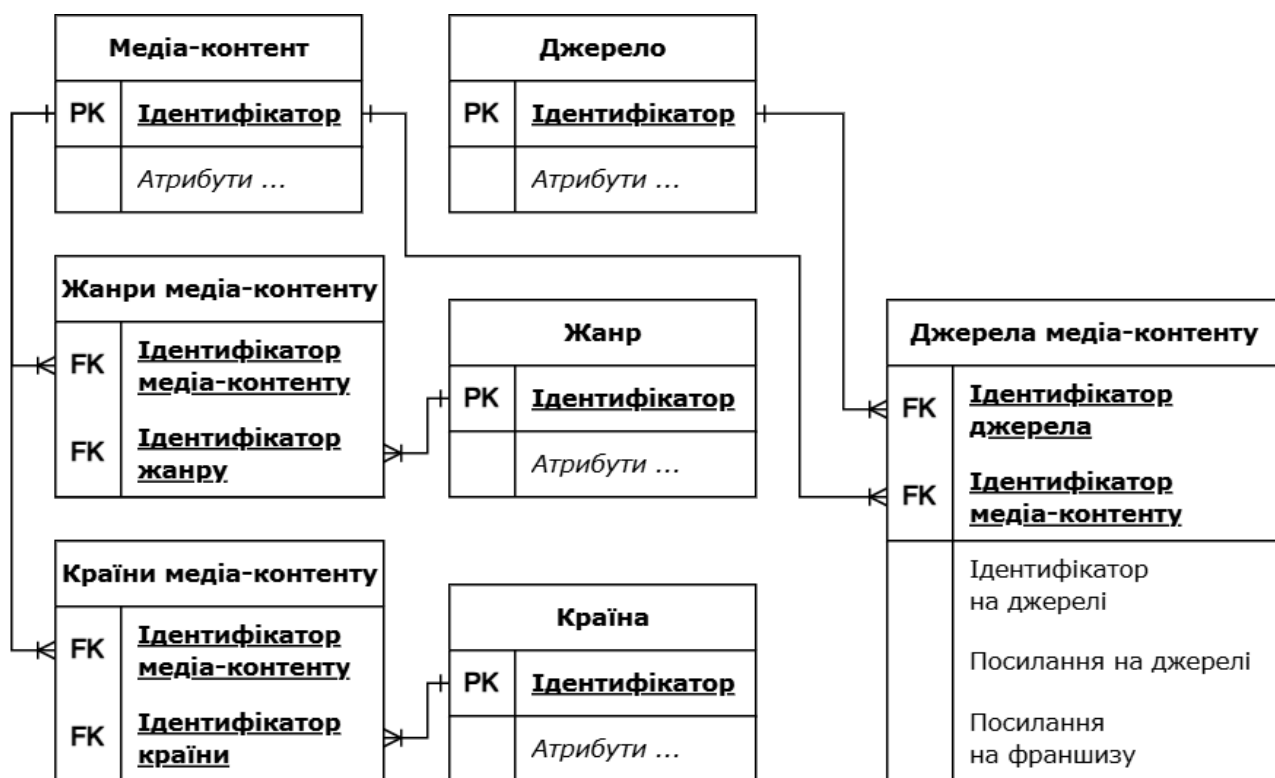


Рисунок 2.9 – Подолання зв’язків «багато до багатьох» сутностей «Медіа-контент», «Джерело», «Жанр» та «Країна»

Аналогічна проблема наявна для сутностей «Одиниця лексики», «Категорія» та «Приклади використання», подолання якої зображене на рисунку 2.10.



Рисунок 2.10 – Подолання зв’язків «багато до багатьох» сутностей «Одиниця лексики», «Категорія» та «Приклади використання»

Тепер введені проміжні таблиці (див. рисунок 2.8 та 2.10) містять зовнішні ключі відповідних таблиць. Після подолання усіх зв’язків «багато до багатьох» проведено етап нормалізації – оптимізації структури БД для усунення надлишкових даних і забезпечення цілісності. Задачами її проведення є забезпечення того, щоб наявна система перебувала у всіх трьох нормальних формах (НФ):

- таблиці проекрованої БД вже знаходяться у першій НФ, оскільки всі атрибути у них атомарні;
- таблиці проекрованої БД вже знаходяться у другій НФ, оскільки всі їхні непервинні атрибути функціонально повно залежать від первинних ключів;
- таблиці проекрованої БД вже знаходяться у третій НФ, оскільки кожен їх непервинний атрибут не є транзитивно залежним від первинного ключа.

На рисунку 2.11 наведено зображення логічної ER-діаграми для лексики:

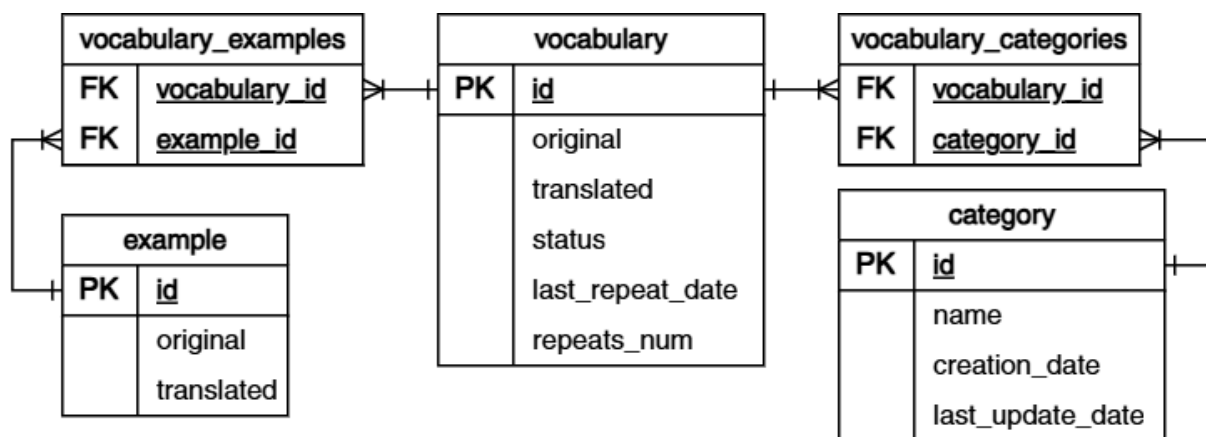


Рисунок 2.11 – Логічна ER-діаграма фрагменту БД інтерактивного застосунку для роботи з лексикою

На рисунку 2.12 наведено зображення логічної ER-діаграми для медіа:

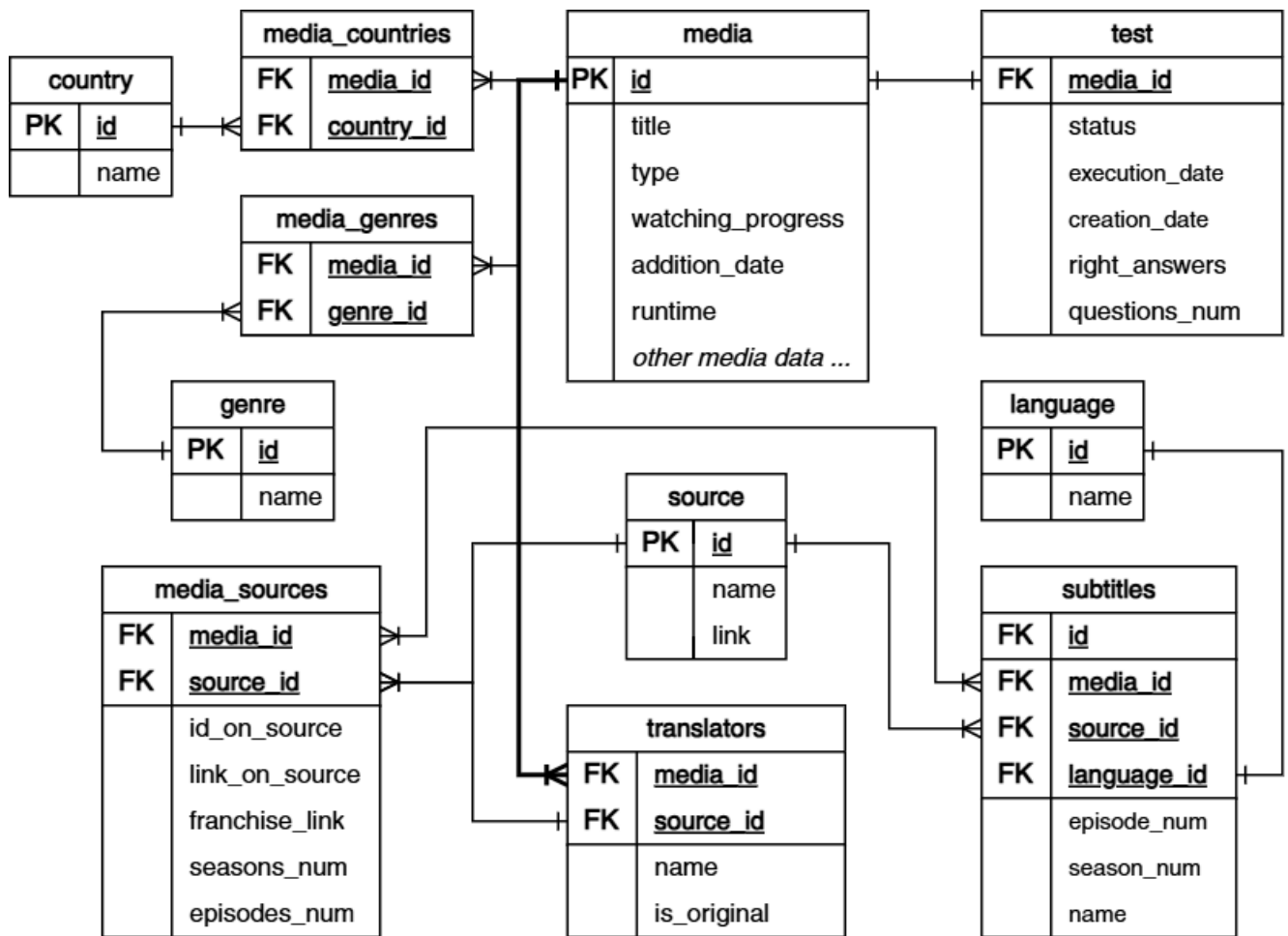


Рисунок 2.12 – Логічна ER-модель фрагменту БД
інтерактивного застосунку для роботи з медіа

Повний перелік атрибутів та первинних ключів сутностей подано у таблицях Е.1-Е.16. Спроектowana БД забезпечує зручне зберігання, швидкий доступ, цілісність даних і можливість ефективного управління та аналізу інформації.

3 РЕАЛІЗАЦІЯ ІНТЕРАКТИВНОГО ЗАСТОСУНКУ ДЛЯ ВИВЧЕННЯ ІНОЗЕМНОЇ МОВИ НА ОСНОВІ МЕДІА-КОНТЕНТУ ІЗ ВИКОРИСТАННЯМ OPENAI

3.1 Програмна реалізація інтерактивного застосунку з використанням OpenAI

Структурна схема програмного продукту реалізовано у вигляді діаграми компонентів UML, що відображає основні програмні модулі, їхні інтерфейси та взаємозв'язки [15]. Застосунок має модульну архітектуру, що забезпечує гнучкість, масштабованість та можливість розширення функціоналу. На рисунку 3.1 зображені модульні складові загального компоненту «Server module».

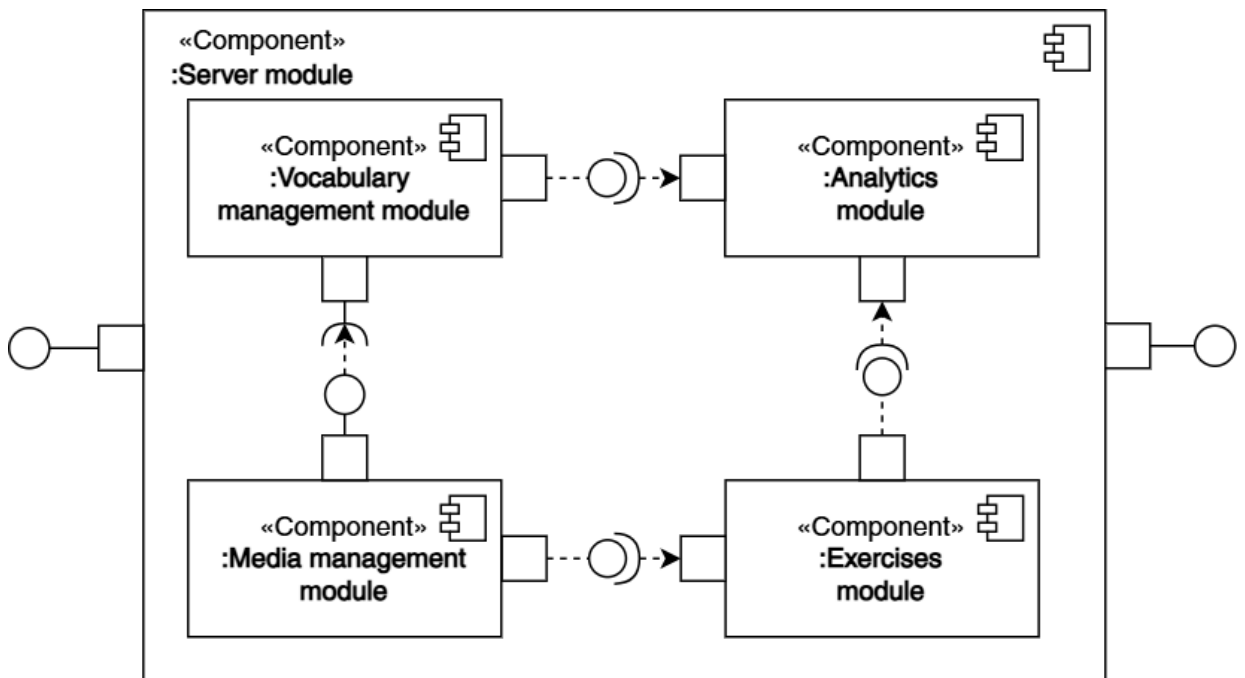


Рисунок 3.1 – UML діаграма компонентів серверної частини «Server module»

З рисунку 3.1 можна побачити такі інформаційні зв'язки [26]:

- «Media Management Module» надає доступ до навчального контенту для «Exercises Module» та «Vocabulary Management Module». Медіа-контент є обов'язковою складовою для генерації аудіо-тестів у модулі вправ та додатковим ресурсом словникового матеріалу для взаємодії із модулем лексики;

- «Exercises Module» зберігає результати тестування. «Vocabulary Management Module» керує даними лексики. Разом вони надають дані для статистики в «Analytics Module»;
 - «Server Module» координує взаємодію між вищевказаними модулями.
- На рисунку 3.2 зображені усі складові застосунку та їх взаємодії.

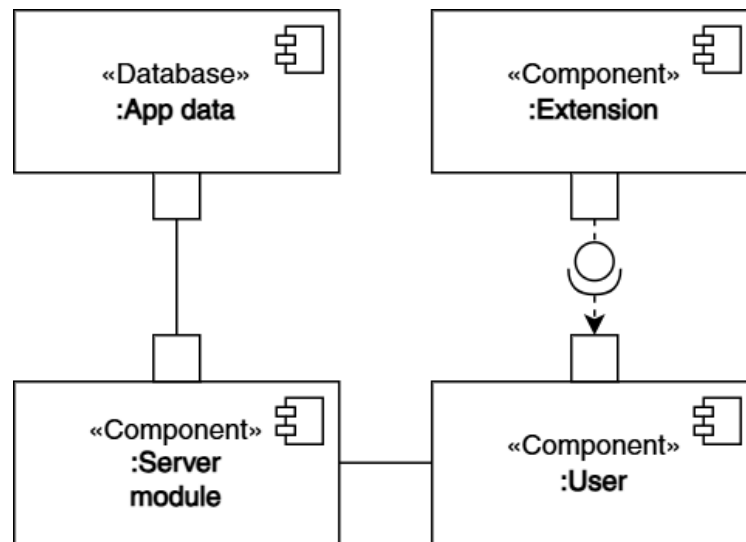


Рисунок 3.2 – UML діаграма компонентів
інтерактивного застосунку

З рисунку 3.2 можна вивести наступні інформаційні зв'язки: «Server Module» звертається до «App Data» для збереження або отримання інформації, а «User Component» може взаємодіяти із «Extension Component» та «Server Module».

Структура програмного комплексу з точки зору класів, їх атрибутів, методів і зв'язків між ними та взаємодія між об'єктами в часовій послідовності зображені на рисунках 2.1 та 2.2 відповідно. Детальний опис роботи кожного з модулів застосунку представлено у підрозділі 2.2.

Основною мовою програмування було обрано Python через широкий спектр готових бібліотек, що значно пришвидшує і спрощує розробку.

Для виконання HTTP-запитів обрано модифікований варіант бібліотеки «requests» – «curl_cffi.requests», що на відміну від оригінальної версії використовує обгортку над libcurl з підтримкою CFFI і дозволяє обробляти сайти з підвищеним захистом від ботів [16]. Такий вибір обумовлений потребою у стабільному та безпечному зборі даних з онлайн-ресурсів.

Для зберігання даних використано легку вбудовану СУБД «SQLite», що є простим й ефективним рішенням без потреби в окремому сервері та малими витратами ресурсів [17]. У ролі засобу її інтеграції з Python використано стандартний модуль «sqlite3», який не потребує ніяких сторонніх залежностей [18].

Для реалізації функціоналу з пошуку синонімів застосовано бібліотеку «`nltk.corpus.wordnet`» [19]. Це дозволить ефективно працювати з синонімічними рядами і покращити варіативність лексики у навчальному процесі.

Для перекладу тексту інтегровано дві моделі з бібліотеки «`deep_translator`» – «`GoogleTranslator`» та «`MyMemoryTranslator`», що дозволяє використовувати більш відповідний інструмент перекладу у конкретній ситуації [20].

Графічний інтерфейс застосунку реалізовано за допомогою «QML (Qt Modeling Language)» – мови опису інтерфейсів із підтримкою реактивності, анімацій і масштабованих компонентів, що надає високий рівень гнучкості системи [21].

Для створення розширення до браузера використано стандартні технології для розробки клієнтських веб-компонентів – «HTML», «CSS» і «JavaScript» – що дозволяють ефективно реалізувати інтерактивні елементи: інтеграцію з медіа-контентом, збір даних із веб-сторінок та взаємодію з API системи.

Зв'язок між клієнтської та серверними частинами організовано за допомогою легкого мікрофреймворку Python – «Flask», який має модульну архітектуру і добре підходить для побудови гнучких серверних рішень [22].

Для генерації інтерактивних тестів у застосунок інтегровано «OpenAI API», а саме модель GPT-4, що демонструє переваги над альтернативними рішеннями у якості генерації тексту, контекстній гнучкості та адаптивності у налаштуванні [23]. Ця модель точно виділяє смислові акценти та логічні зв'язки між твердженнями, що забезпечує релевантність, точність і змістовність тестових запитань, а збереження широкого контексту сприяє формуванню питань вищого рівня складності.

Початковим етапом реалізації логіки системи є автоматизований пошук медіа-даних, джерелом яких обрано IMDb [24]. Оскільки застосунок працює локально, з метою доступу до великого обсягу актуальної інформації без перевантаження пам'яті пристрою використовується динамічна інтеграція через надсилання запитів до офіційного GraphQL-інтерфейсу IMDb. Це дозволяє здійснювати пошук

у режимі реального часу без потреби зберігати зайві дані на диску. У межах цієї логіки реалізовано функцію-обгортку для формування запитів й обробки отриманих метаданих, лістинг якої подано на рисунку Ж.1.

З огляду на потенційне розширення кількості підтримуваних джерел субтитрів і супутніх метаданих, у системі передбачено можливість взаємодії із різними сервісами медіа-контенту. Для цього застосовано шаблон проєктування «Фабричний метод» («Factory method»), архітектура якого базується на абстрактному класі, що визначає спільний інтерфейс для всіх парсерів та перелік методів, обов'язкових до реалізації у класах-нащадках [25]. Це забезпечує інкапсуляцію логіки, зниження зв'язності компонентів та дотримання принципу відкритості/закритості (SOLID), залишаючи систему відкритою до розширення, але закритою до модифікації. Лістинг представлено на рисунку Ж.2.

З метою перетворення субтитрів на зв'язний текст для його подальшої інтерпретації цільовою мовою, вони проходять попередню обробку: видаляються службові мітки (номери рядків, часові позначки, HTML-теги) та зайві пропуски. Функція перекладу (рисунок Ж.3) підтримує багатопоточність, що дозволяє уникнути перевантаження системи та забезпечити швидкість генерації тексту. Максимальна кількість одночасних потоків визначається динамічно на основі доступних ядер процесора. Підтримувані перекладачі: «GoogleTranslator» та «MyMemoryTranslator». Також реалізовано механізм додавання префіксу, що створює певний керований контекст для підвищення точності машинного перекладу в окремих випадках – після перекладу цей префікс вилучається.

Наступним кроком є забезпечення мовної варіативності та підвищення семантичної насиченості тексту. З цією метою у системі реалізовано функцію отримання синонімів для заданого слова, яка автоматично усуває повтори, а також виключає із результату вихідне слово, забезпечуючи релевантність отриманих синонімів (рисунок Ж.4). Повернені варіанти здатні підвищити лінгвістичну якість системи.

Для забезпечення повноцінної взаємодії між клієнтською частиною браузерного розширення та серверною логікою було реалізовано набір REST-інтерфейсів, які обслуговують основні сценарії використання системи, зокрема створено функції: отримання доступних субтитрів обраного медіа (рисунок Ж.5), збереження

словникових одиниць (рисунки Ж.6 та Ж.7), фіксації поточної часової мітки (рисунки Ж.8) та автоматичного перекладу переданого тексту із опціональної можливістю отримання синонімів (рисунок Ж.9).

Функціональність розширення активно використовує API бекенду для динамічного отримання й оновлення інформації. При увімкненні розширення повертається список субтитрів, прив'язаних до поточного відео, після чого їх контейнер вбудовуються у DOM плеєра (рисунки Ж.10, Ж.11). Система реалізує двокомпонентне відстеження часу: кожні 200 мс зчитується час відтворення з localStorage для синхронізації субтитрів та щосекунди виконується оновлення прогресу перегляду на сервері (рисунок Ж.12). Для додавання лексики у словник, спершу завантажуються вже створені категорії, що поліпшує досвід організації навчального матеріалу (рисунок Ж.13). Користувач може інтерактивно взаємодіяти із текстовими доріжками: зберігати їх у персональний словник, переглядати приклади вживання з відео, отримувати автопереклад й підказки до слів і речень, а також синоніми у випадку їх наявності (рисунки Ж.14, Ж.15). Створюється візуальний індикатор статусу активності розширення, а при перемиканні стану усе оновлюється в реальному часі без перезавантаження сторінки (рисунок Ж.16).

Функція для створення тестів (рисунок Ж.17) звертається до OpenAI API для автоматичної генерації тестових запитань на основі фрагмента субтитрів з переглянутого користувачем медіа-контенту. У запиті враховуються тип (наприклад, фільм, серіал) і назва контенту, що дає змогу моделі покращити релевантність та точність формування запитань і відповідей до них.

3.2 Інтерфейс користувача інтерактивного застосунку

При проєктуванні інтерфейсу основною метою стало створення інтуїтивно зрозумілої та логічної системи взаємодії, що передбачає мінімізацію кількості кроків до ключових функцій програми та обробку потенційних помилок користувача.

Графічна оболонка застосунку має чітку структуру, представлену в головному меню додатку, що охоплює всі основні функціональні блоки системи: медіа-

бібліотеку, словники, вправи (навчальні картки та тестування), аналітику й персоналізацію (розширення та управління резервними даними).

Розділ медіа-бібліотеки (рисунок 3.3) слугує центральним простором для взаємодії користувача з медіа-контентом, що використовується у навчальному процесі.

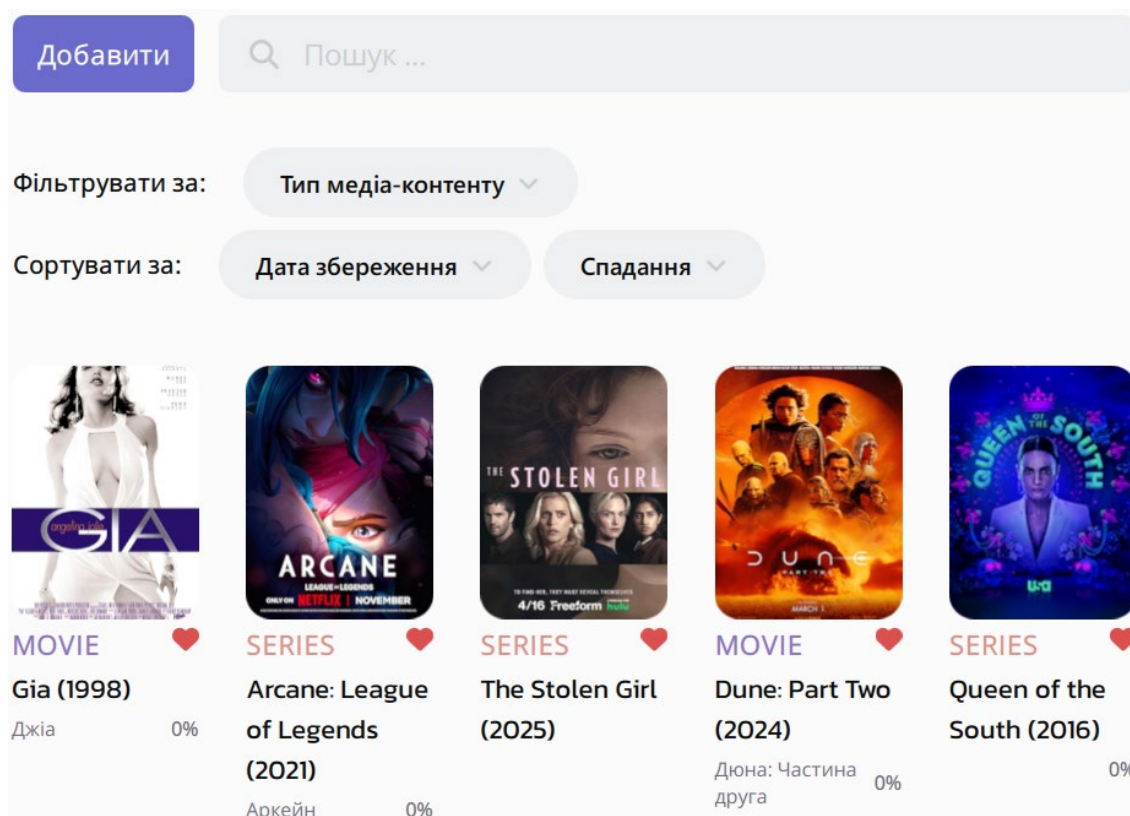


Рисунок 3.3 – Медіа-бібліотека

Сторінка «Словники» (рисунок 3.4) призначена для керування й систематизації особистих словникових наборів: їх створення, сортування та пошук.

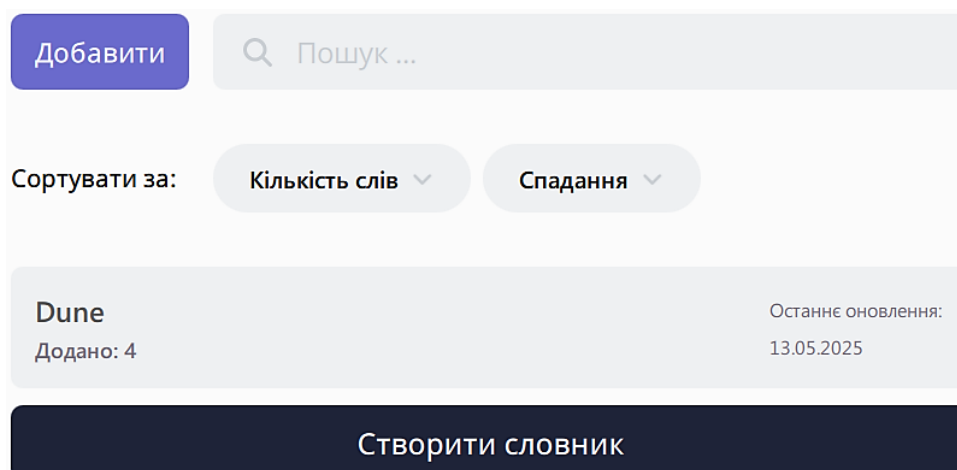


Рисунок 3.4 – Сторінка «Словники»

Секція «Вправи» (рисунки 3.5) слугує для закріплення матеріалу та включає інтервальне повторення (реалізовано у форматі карток для гортання) та тестування (дозволяє створити новий тест на основі переглянутого медіа або пройти збережений).

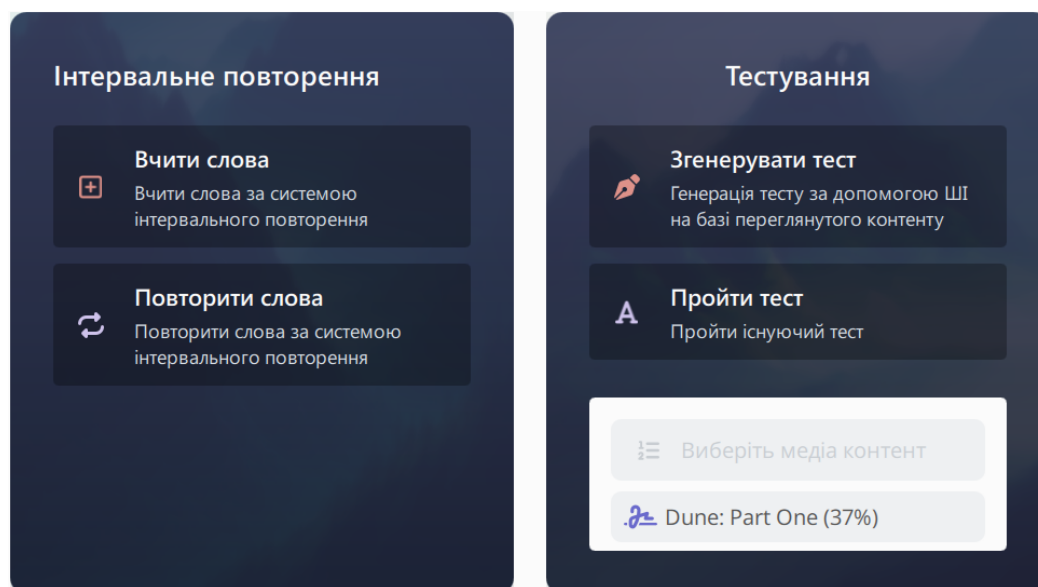


Рисунок 3.5 – Секція «Вправи»

У розділі аналітика зображені діаграми, що допомагають відстежувати прогрес у вивченні лексики та проходженні тестів (рисунки 3.6).

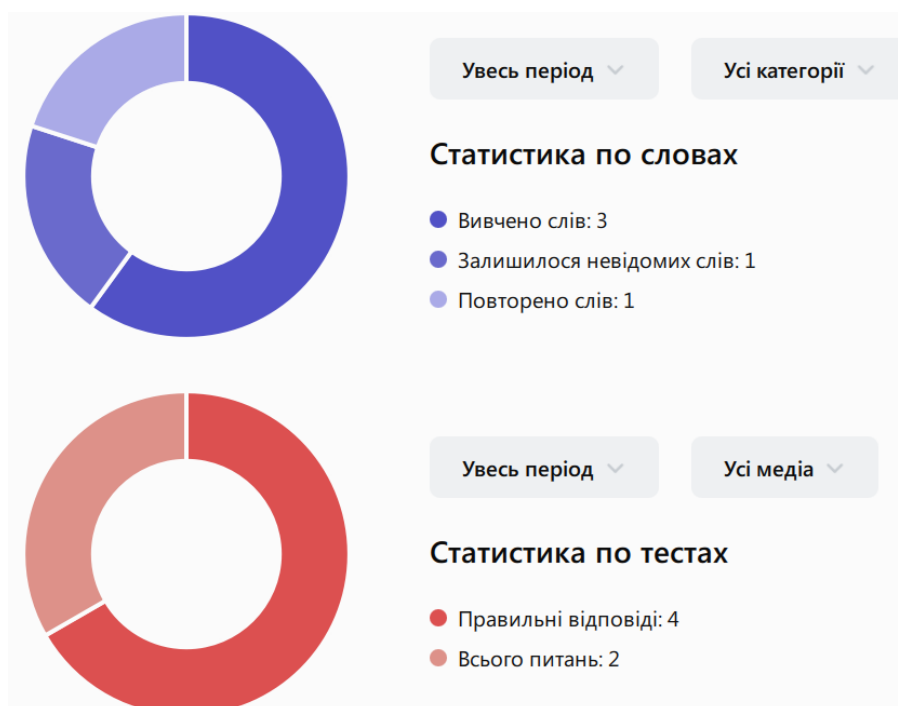


Рисунок 3.6 – Розділ «Аналітика»

Користувач може обрати часовий період для перегляду статистики (сьогодні, тиждень, за місяць, увесь час). Також результати статистики по словах можна фільтрувати за категоріями словників, а дані по тестуванню за медіа-контентом.

У розділі «Персоналізація» (рисунок 3.7) користувачеві надаються інструменти для створення й завантаження резервної копії та доступу до браузерного розширення для інтерактивної взаємодії із субтитрами.

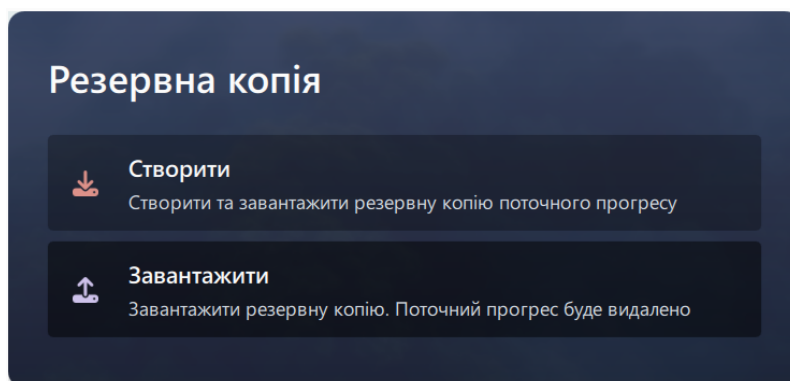


Рисунок 3.7 – Розділ «Персоналізація» – управління даними

Розширення є інтерактивним інструментом, який доповнює функціональність основного застосунку безпосередньо під час перегляду відеоконтенту на зовнішніх платформах. Воно включає індикатор активності – піктограму в панелі розширень, що змінює свій вигляд залежно від стану активності (рисунок 3.8).

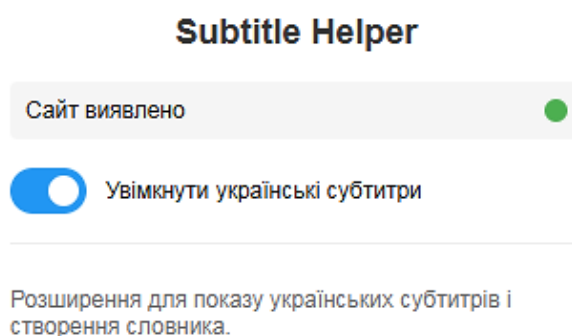


Рисунок 3.8 – Індикатор активності розширення

Панель керування надає користувачу доступ до базових налаштувань для гнучкого керування відображенням субтитрів (рисунок 3.9). Їх можна вмикати, вимикати субтитри та обирати режим перегляду – лише оригінал чи переклад або обидва одночасно для кращого розуміння.

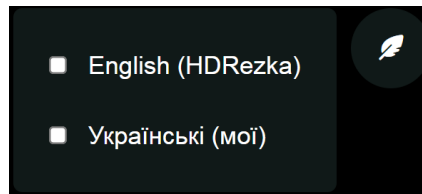


Рисунок 3.9 – Панель керування субтитрами

Інтерактивність субтитрів проявляється при виділенні тексту – одразу з’являється спливаюче вікно з перекладом і синонімами (за їх наявності) для обраного фрагменту (рисунок 3.10).

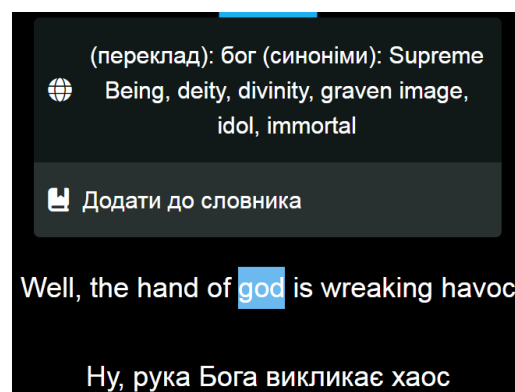


Рисунок 3.10 – Вікно підказок із перекладом та синонімами

Виділені фрагменти тексту можна додати до словника скориставшись повноцінним набором функцій, як й у самому застосунку (рисунок 3.11).

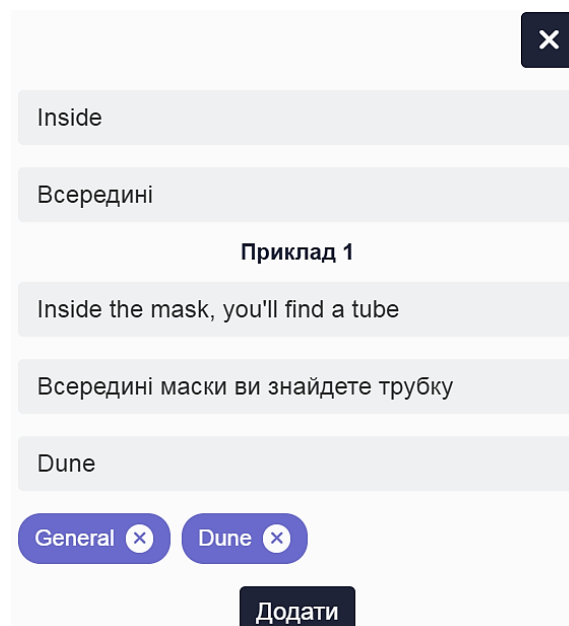


Рисунок 3.11 – Вікно додавання одиниці лексики у розширенні

Для зручності та кращого запам'ятовування автоматично додається приклад використання – речення, з якого взято слово. Переклад лексики та прикладу також генерується автоматично. При введенні назви словника користувач отримує динамічні підказки з уже створених категорій у БД.

Ще одним важливим аспектом є режим фіксації субтитрів – під час взаємодії користувача із текстом (наприклад, при натисканні чи виділенні) автоматичне оновлення субтитрів тимчасово призупиняються, щоб уникнути ризику втрати уваги через зміну інформації та не перешкоджати навчанню й обробці матеріалу.

3.3 Тестування інтерактивного застосунку для вивчення іноземної мови

Під час проведення тестування функціональних блоків було виявлено проблему повільної обробки великих обсягів тексту при перекладі, а також обмеження API використаних бібліотек на кількість символів. Це спричиняло неповні переклади. Рішенням стала реалізація багатопотокової системи перекладу з обробкою фрагментів у черзі, що забезпечило стабільну обробку без помилок.

Іншим складним етапом реалізації стала інтеграція інтерактивних елементів через браузерне розширення без порушення структури та стилів DOM-елементів стрімінгових платформ. Динамічне завантаження контенту ускладнювало цю задачу. Рішенням стало використання гнучкої архітектури розширення з ізольованими стилями, відкладеної ініціалізації елементів та контейнерного підходу.

Особливу увагу приділено формуванню запитів до OpenAI API. Для отримання уніфікованого шаблону відповіді запити були ретельно структуровані: уточнено інструкції, мінімізовано кількість вхідних токенів та забезпечено чітке розуміння моделлю завдання й очікуваного формату відповіді. Це дозволило гарантувати стабільність генерації результатів та ефективність їх інтеграції в систему.

Також було протестовано кожен функціональний компонент інтерфейсу застосунку. Основні функціональні можливості модуля роботи із медіа-бібліотеки включають: керування вмістом, сортування, фільтрацію, локальний (в межах особистої бібліотеки) та глобальний (забезпечує доступ до зовнішньої БД IMDb, рисунок 3.12) пошуки.

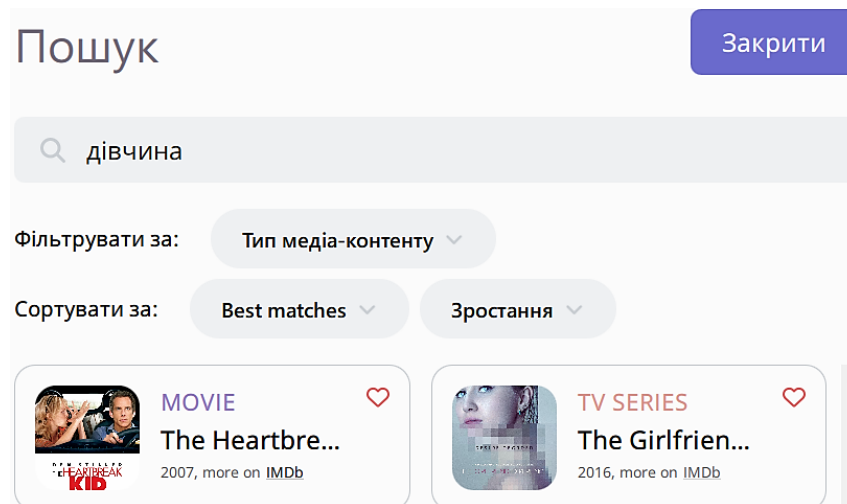


Рисунок 3.12 – Сторінка медіа-пошуку під'єднаного до БД IMDb

При кліку на об'єкт медіа відкривається його окрема сторінка із розширеною інформацією та інструментами для подальшої взаємодії (рисунок 3.13).

The VVitch: A New-England Folktales

The Witch Відьма: Нова англійська казка

7/10 · 2016 · 1h 32m · Movie · United States, Canada, United Kingdom

Драма Таємниця Трилер Фентезі Жахи

Про що «The VVitch: A New-England Folktales»

Відьма: Режисер Роберт Еггерс. З Анею Тейлор-Джой, Ральфом Інесоном, Кейт Дікі, Харві Скрімшоу. Ізольована пуританська сім'я в Новій Англії 1630-х років розгромлена силами чаклунства та одержимості.

Рисунок 3.13 – Сторінка медіа-контенту

Сторінка виступає інформаційним центром відео, надаючи ключові мета-дані та необхідні функції для навчання: наявна інтеграція із зовнішніми сервісами перегляду (рисунок 3.14) і можливість генерування перекладу субтитрів (рисунок 3.15).



Рисунок 3.14 – Інтеграція із зовнішніми сервісами перегляду

Субтитри



Згенерувати субтитри

Мова	База
Англійська	English (HDRezka)
Українська	Українські (мої)
Метод	Назва
GoogleTranslate	English (HDRezka)
OpenAI	Українські (мої)
Згенерувати	

Рисунок 3.15 – Інтеграція із зовнішніми сервісами перегляду

При переході до конкретної категорії у модулі «Словники» відкривається детальна сторінка зі статистикою навчання (рисунок 3.16) та можливістю додавати слова й приклади.

Dune

4

- Нові слова: 1
- Слова для повторення: 1
- Вивчені слова: 2

Словник: Dune

Додати

Сортувати за: Оригінал Зростання

cat
cat

Вивчається
Повтерено: 1

cut
cut

Нове
Повтерено: 0

Закрити

Рисунок 3.16 – Детальна сторінка словника

У модулі «Вправи» інтервальне повторення реалізовано у форматі карток для гортання (рисунок 3.17) та тестування – можна створити новий тест на основі переглянутого медіа або пройти вже збережений (рисунок 3.18).

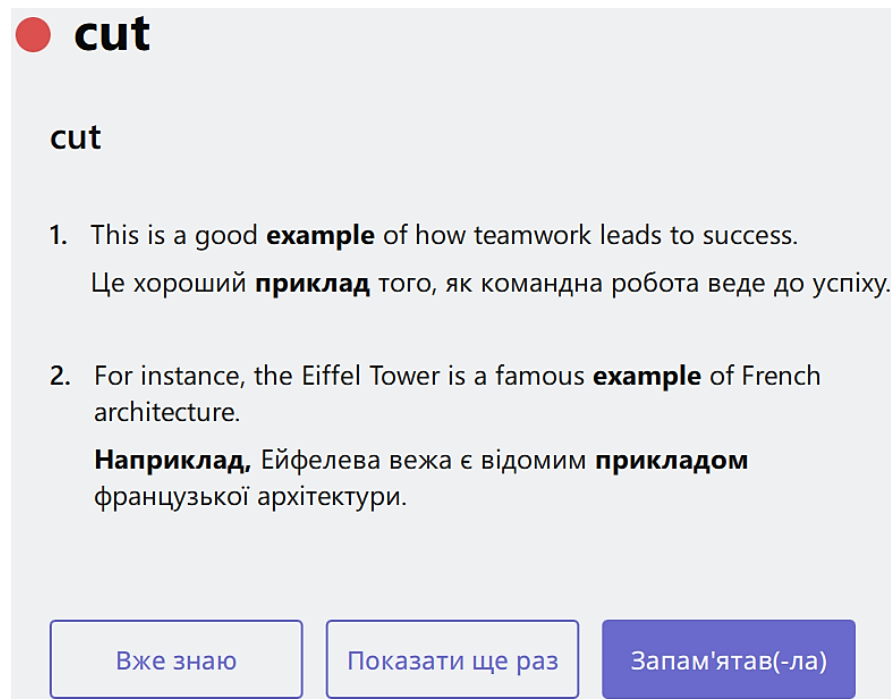


Рисунок 3.17 – Секція інтервально повторення лексики

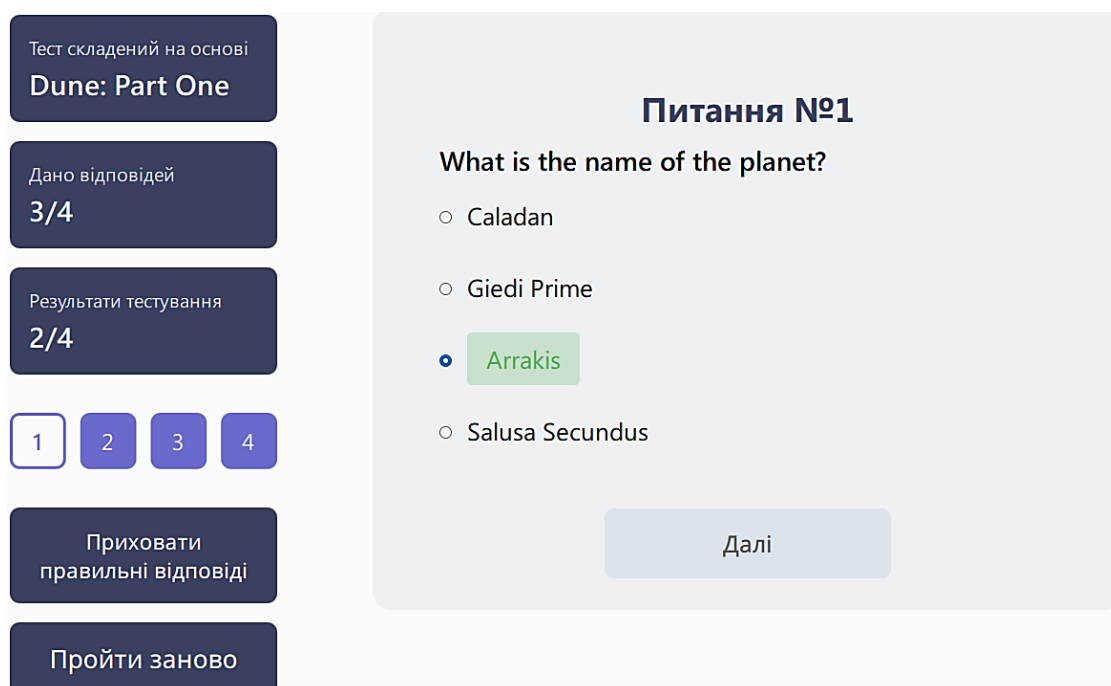


Рисунок 3.18 – Секція тестування

Проведене тестування дозволило переконатися у відповідності застосунка заданим вимогам та зручності взаємодії у типових сценаріях використання.

ВИСНОВКИ

Під час виконання кваліфікаційної роботи було:

1. Проведено аналіз предметної області для виявлення типових труднощів користувачів у процесі вивчення мови через медіа-контент.
2. Здійснено огляд існуючих аналогів, виявлено їхні переваги та недоліки, що дозволило сформулювати вимоги до застосунку для вивчення англійської мови.
3. Описано структуру інтерактивного застосунку, включаючи основні функціональні модулі та схему зберігання й обробки навчального матеріалу.
4. Розроблено алгоритмічне та інформаційне забезпечення застосунку.
5. Реалізовано інтерфейс та функціонал застосунку, що включає можливість перегляду медіа з інтерактивними елементами навчання; особливу увагу приділено зручності інтерфейсу і стабільності його взаємодії з DOM-структурами стрімінгових сервісів.
6. Інтегровано дієві інструменти вивчення мови, такі як субтитри, повторення, переклад, визначення нових слів та виразів в контексті; розробка здійснювалася з використанням бібліотек машинного навчання (зокрема для перекладу) та багатопоточних й асинхронних механізмів для ефективної роботи з великими обсягами текстових даних.
7. Виконано тестування різних форм інтерактивності; виявлено та усунуто помилки, пов'язані з перекладом довгих субтитрів та інтеграцією із зовнішніми платформами.
8. Реалізовано багатофункціональний застосунок, поєднує перегляд цільового відео-контенту у зручному форматі, інтерактивні методи взаємодії із субтитрами, вправи для навчання та інструменти на основі OpenAI, що дозволяє вивчати лексику безпосередньо в контексті використання, зберігати обрані слова і вирази у тематичні словники для подальшого систематичного опрацювання, а також закріплювати отримані знання за допомогою вправ. Застосунок доповнено браузерним розширенням для інтеграції з платформами перегляду.
9. Застосунок може бути особливо корисним для індивідуального навчання, оскільки дає змогу органічно поєднувати навчальний процес із дозвіллям,

перетворюючи перегляд медіа-контенту на ефективну мовну практику. Він сприяє формуванню стійкої мотивації до навчання, дозволяючи користувачеві вивчати мову у комфортному ритмі та у звичному для себе середовищі.

10. У майбутньому доцільно доповнити функціонал застосунку за рахунок розширення типів вправ, покращення аналітики навчального прогресу, забезпечення можливості роботи із локальними відео-файлами та файлами субтитрів, а також інтеграції із ширшим спектром стримінгових сервісів, включно з тими, які початково не передбачають наявності субтитрів.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

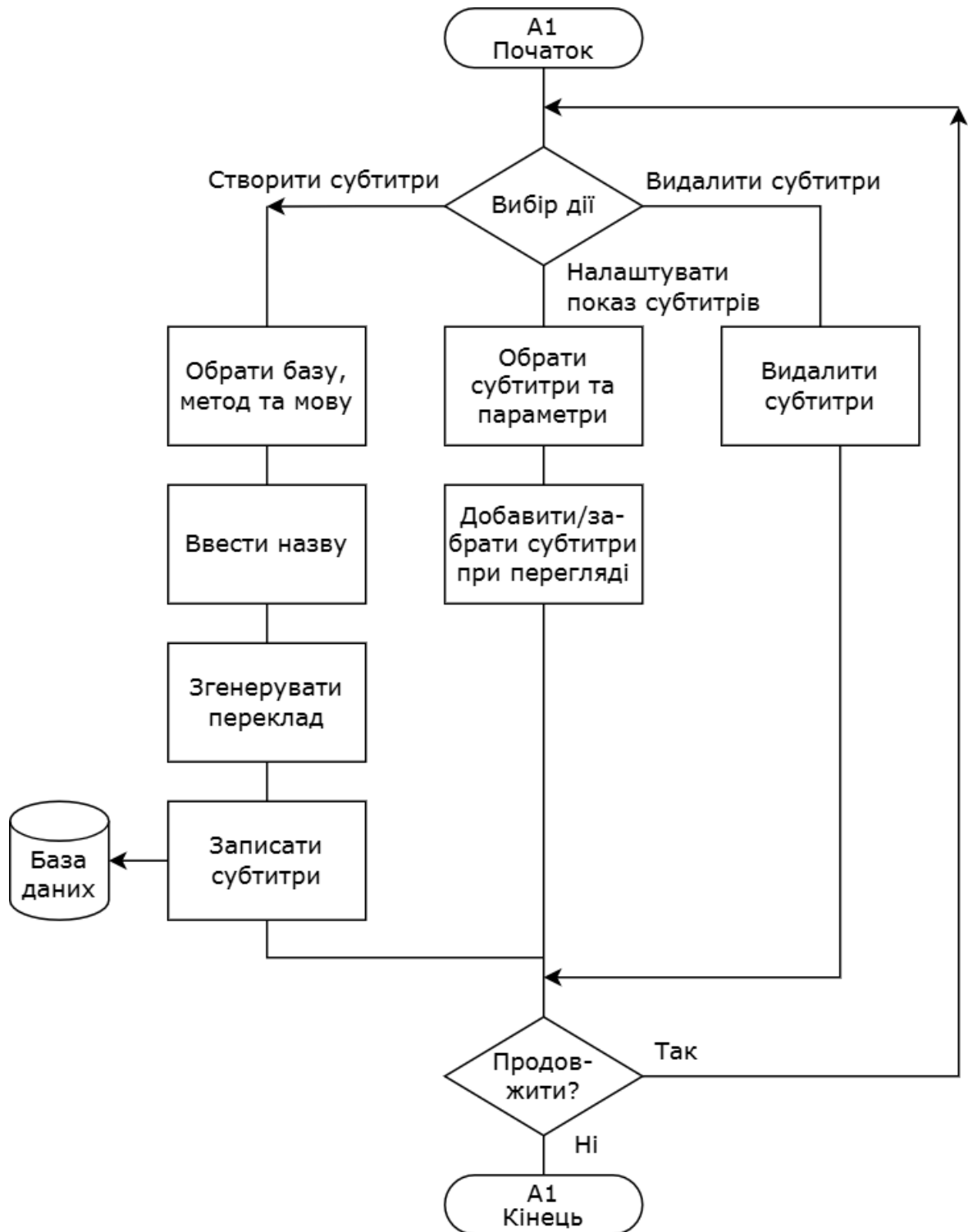
1. Kozanitis A., Nenciovici L. Effect of active learning versus traditional lecturing on the learning achievement of college students in humanities and social sciences: a meta-analysis. *Higher Education*. 2023. Vol.86(6). P.1201-1226.
2. Balenović K., Proroković J. Foreign language vocabulary development: media-driven learning in the informal context. *Suvremena lingvistika*. Vol.49(96). 2023. P.175-201.
3. Hayitova N. Y. Enhancing listening skills through movies: an effective language learning strategy. *Ethiopian International Journal of Multidisciplinary Research*. 2025. Vol.12(2). P.256-259.
4. Ghanizadeh A., Tabeie M., Pourtousi Z. The role of university instructor's narrative in students' sustained attention, emotional involvement and cognitive learning. *Journal of Applied Research in Higher Education*. 2024. Vol.16(1). P.195-207.
5. Short-Form Videos Degrade Our Capacity to Retain Intentions: Effect of Context Switching on Prospective Memory. URL: <https://dl.acm.org/doi/10.1145/3544548.3580778>.
6. Green M.C., Brock T.C. The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*. 2000. Vol.79(5). P.701-721.
7. Hien L.T., Phuong N.T. The effectiveness of the storytelling technique on students' achievement and motivation in English speaking skills. *Journal of Applied Business and Economics*. 2023. Vol.10(1). P.1-15.
8. Nalyvaiko O., Nuralieva L., Syvolotska L. Immersion in a foreign environment as an effective way of development of communicative competence in the process of learning foreign languages. *Scientific Notes of the Pedagogical Department*. 2021. Vol.49(5). P.42-51.
9. Sulistyawati W., Sumartiningsih I., Priatiningsih S. The Influence of Interactive Learning Media to Increase Learning Motivation and Learning Outcomes for Early Childhood. *Journal of Education Technology and Innovation*. 2024. Vol.7(1). P. 38-46.

10. University Students' Perception of Learning Listening Comprehension with Interactive Multimedia Platforms: Moodle and Thinglink-Based Approach. URL: <https://eudl.eu/doi/10.4108/eai.24-10-2023.2342125>.
11. Cake. URL: <https://cake.day/>.
12. Listening. URL: <https://learnenglish.britishcouncil.org/skills/listening>.
13. Діаграма послідовності (sequence diagrams). URL: <https://www.maxzsim.com/sequence-diagrams/>.
14. Для чого потрібні UML діаграми? URL: <https://foxminded.ua/uml-diagramy/>.
15. Ermel G., Farias K., Gonçalves L.J., Bischoff V. Supporting the Composition of UML Component Diagrams. *Proceedings of the 14th Brazilian Symposium on Information Systems (SBSI'18)*. 2018. Vol.1(1). P.1-8.
16. curl_cffi. URL: <https://pypi.org/project/curl-cffi/>.
17. An Overview of SQLite: Features, Uses, and Benefits. URL: <https://platodata.network/platowire/an-overview-of-sqlite-features-uses-and-benefits/>.
18. sqlite3 – DB-API 2.0 interface for SQLite databases. URL: <https://docs.python.org/3/library/sqlite3.html>.
19. Sample usage for wordnet. URL: <https://www.nltk.org/howto/wordnet.html>.
20. deep-translator. URL: <https://pypi.org/project/deep-translator/>.
21. QML Applications. URL: <https://doc.qt.io/qt-6.7/qmlapplications.html>.
22. Flask. URL: <https://flask.palletsprojects.com/en/stable/>.
23. The 55 Best AI Tools for 2025 (Tried and Tested). URL: <https://www.synthesia.io/post/ai-tools>.
24. IMDb. URL: <https://www.imdb.com/>.
25. M. Chochlík. Implementing the factory pattern with the help of Reflection. *Computing and Informatics*. 2015. Vol.35(3). P.1001-1020.
26. Климчук А.Б., Турченко І.В. Інтерактивний застосунок для вивчення англійської мови на основі медіа-контенту із використанням OpenAI. *Інтелектуальні інформаційні технології в прикладних дослідженнях» (ІІТАР-2025)*: збірник тез доповідей студентської науково-практичної конференції (Тернопіль, 27-29 травня 2025 р). Тернопіль: ЗУНУ, 2025. С.168-170.

27. Комар М.П., Саченко А.О., Васильків Н.М., Гладій Г.М., Коваль В.С., Ліп'яніна-Гончаренко Х.В. Методичні рекомендації до виконання кваліфікаційної роботи з освітньо-професійної програми «Комп'ютерні науки» спеціальності 122 «Комп'ютерні науки» за першим (бакалаврським) рівнем вищої освіти. Тернопіль: ЗУНУ, 2024. 52 с.

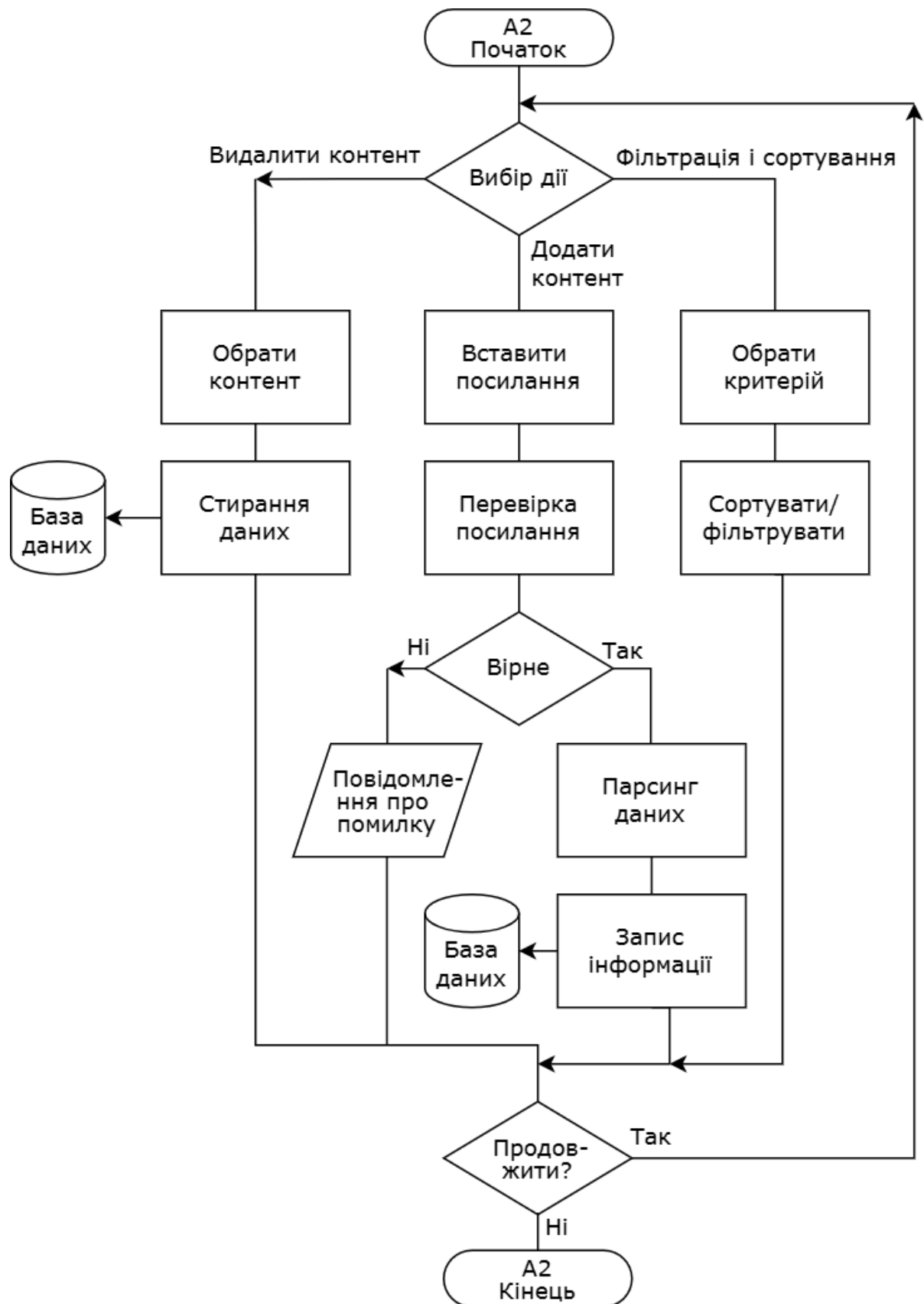
Додаток А

Схема алгоритму роботи із субтитрами



Додаток Б

Схема алгоритму роботи із бібліотекою медіа-контенту



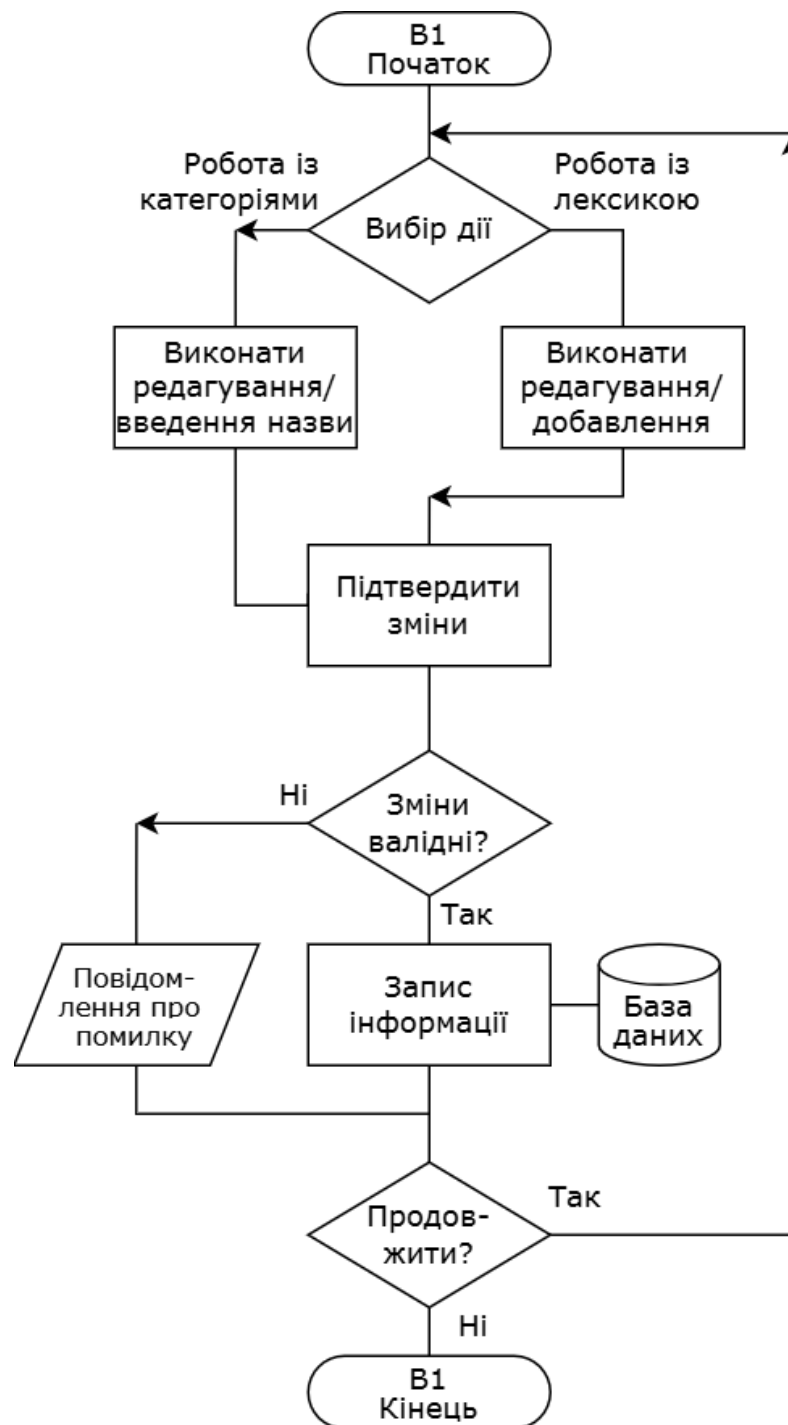
Додаток В

Схема алгоритму інтерактивної взаємодії із субтитрами



Додаток Г

Схема алгоритму роботи із бібліотекою лексики



Додаток Д

Схема алгоритму вивчення лексики



Додаток Е

Атрибути і первинні ключі сутностей БД

Таблиця Е.1 – Атрибути і первинний ключ сутності «Джерело»

Сутність	Атрибути	Первинний ключ
Джерело	Ідентифікатор джерела Назва Посилання на головну	Ідентифікатор джерела

Таблиця Е.2 – Атрибути і первинний ключ сутності «Медіа»

Сутність	Атрибути	Первинний ключ
Медіа	Ідентифікатор медіа Назва Тип Дата збереження Прогрес перегляду Тривалість Рейтинг Рік виходу Опис Обкладинка постеру Посилання на трейлер	Ідентифікатор медіа

Таблиця Е.3 – Атрибути і первинний ключ сутності «Джерела медіа»

Сутність	Атрибути	Первинний ключ
Джерела медіа	Ідентифікатор медіа Ідентифікатор джерела Ідентифікатор на ресурсі Посилання медіа на ресурсі Кількість сезонів Кількість епізодів	Ідентифікатор медіа + Ідентифікатор джерела

Таблиця Е.4 – Атрибути і первинний ключ сутності «Жанр»

Сутність	Атрибути	Первинний ключ
Жанр	Ідентифікатор жанру Назва	Ідентифікатор жанру

Таблиця Е.5 – Атрибути і первинний ключ сутності «Жанри медіа»

Сутність	Атрибути	Первинний ключ
Жанри медіа	Ідентифікатор медіа Ідентифікатор жанру	Ідентифікатор медіа + Ідентифікатор жанру

Таблиця Е.6 – Атрибути і первинний ключ сутності «Країна»

Сутність	Атрибути	Первинний ключ
Країна	Ідентифікатор країни Назва	Ідентифікатор країни

Таблиця Е.7 – Атрибути і первинний ключ сутності «Країни медіа»

Сутність	Атрибути	Первинний ключ
Країни медіа	Ідентифікатор медіа Ідентифікатор країни	Ідентифікатор медіа + Ідентифікатор країни

Таблиця Е.8 – Атрибути і первинний ключ сутності «Звукова доріжка»

Сутність	Атрибути	Первинний ключ
Звукова доріжка	Ідентифікатор медіа Ідентифікатор джерела Ідентифікатор звукової доріжки на джерелі Назва Чи оригінал	Ідентифікатор медіа + Ідентифікатор джерела + Ідентифікатор звукової доріжки на джерелі

Таблиця Е.9 – Атрибути, первинний сутності «Мова»

Сутність	Атрибути	Первинний ключ
Мова	Ідентифікатор мови Назва	Ідентифікатор мови

Таблиця Е.10 – Атрибути і первинний ключ сутності «Субтитри»

Сутність	Атрибути	Первинний ключ
Субтитри	Ідентифікатор субтитрів Ідентифікатор медіа Ідентифікатор джерела Ідентифікатор мови Назва Номер сезону Номер епізоду	Ідентифікатор субтитрів

Таблиця Е.11 – Атрибути і первинний ключ сутності «Тест»

Сутність	Атрибути	Первинний ключ
Тест	Ідентифікатор тесту Ідентифікатор медіа Дата створення Дата проходження Статус Кількість питань Кількість вірних відповідей	Ідентифікатор тесту

Таблиця Е.12 – Атрибути і первинний ключ сутності «Одиниця лексики»

Сутність	Атрибути	Первинний ключ
Одиниця лексики	Ідентифікатор одиниці лексики Оригінал Переклад Статус вивчення Дата останнього повторення Кількість повторень	Ідентифікатор одиниці лексики

Таблиця Е.13 – Атрибути і первинний ключ сутності «Словник»

Сутність	Атрибути	Первинний ключ
Словник	Ідентифікатор словника Назва Дата створення Дата останнього оновлення	Ідентифікатор словника

Таблиця Е.14 – Атрибути і первинний ключ сутності «Лексика словника»

Сутність	Атрибути	Первинний ключ
Лексика словника	Ідентифікатор лексики Ідентифікатор словника	Ідентифікатор лексики + Ідентифікатор словника

Таблиця Е.15 – Атрибути і первинний ключ сутності «Приклади використання»

Сутність	Атрибути	Первинний ключ
Приклади використання	Ідентифікатор прикладу Оригінал Переклад	Ідентифікатор прикладу

Таблиця Е.16 – Атрибути і первинний ключ сутності «Приклади використання лексики»

Сутність	Атрибути	Первинний ключ
Приклади використання лексики	Ідентифікатор лексики Ідентифікатор прикладу	Ідентифікатор лексики + Ідентифікатор прикладу

Програмний код основного функціоналу застосунку

```

1 def media_search(text: str, numResults: int = 21) -> Media | None:
2     """ Get media content by text request using IMDB search """
3     variables_template, headers, params = ..., {...}, {...}
4     response = requests_parser.make_request(
5         url="https://caching.graphql.imdb.com/", headers=headers, params=params)
6
7     data, media = response.json(), []
8     data = data["data"]["results"]["edges"]
9
10    for entity in data:
11        # get data ...
12        media.append(Media(
13            imdb_id = imdb_id, titles = Title(title_original = title),
14            poster = poster, release_year = release_year, media_type = media_type))
15    return media

```

Рисунок Ж.1 – Лістинг функції опрацювання метаданих IMDb

```

1 class SourceParser(ABC):
2     def __init__(self, link: str, imdb_id: str):
3         self.__imdb_id, self.__link, self.__soup = imdb_id, link, None
4
5     @abstractmethod
6     def _get_franchise(self) -> str | None:
7         """_get_franchise"""
8         pass
9
10    @abstractmethod
11    def _is_correct_genre(self) -> bool:
12        """_is_correct_genre"""
13        pass
14
15    @abstractmethod
16    def _get_media_source_id(self) -> int:
17        """_get_media_source_id"""
18        pass
19
20    @abstractmethod
21    def _get_translators(self) -> list[Translator]:
22        """_get_translators"""
23        pass
24
25    @abstractmethod
26    def _get_seasons_episodes_number(self) -> dict[str, int | None]:
27        """_get_seasons_episodes_number"""
28        pass
29
30    def parser(self) -> MediaSource | None:
31        """Parse content from the link"""
32        response = requests_parser.make_request(url=self.__link).content
33        self.__soup = BeautifulSoup(response, "html.parser")
34        if not self._is_correct_genre(): return None
35        s_e_num = self._get_seasons_episodes_number()
36
37        return MediaSource(
38            imdb_id=self.__imdb_id, source_id=SOURCES["hdrezka"].id,
39            id_on_source=self._get_media_source_id(), link_on_source=self.__link,
40            franchise_link=self._get_franchise(), translators=self._get_translators(),
41            seasons_number=s_e_num["seasons"], episodes_number=s_e_num["episodes"],
42            subtitles=self._get_subtitles())

```

Рисунок Ж.2 – Лістинг класу-інтерфейсу взаємодії сервісів медіа

```

1 def generate_subtitles(
2     target_language: Language,
3     orig_base: Subtitles,
4     method: str, name: str
5 ) -> Subtitles | None:
6     """ Generate subtitles based on the original subtitles """
7     orig_subs = extract_subtitles(subtitles_id=orig_base.id)
8
9     from src.backend.database.manager import Manager as MediaManager
10    media_manager = MediaManager()
11    source_lang = media_manager.get_language_by_id(orig_base.language.id)
12
13    orig_plain_subs = extract_text_from_webvtt(webvtt_content=orig_subs)
14
15    if method.lower() == "googletranslate":
16        target_lang_code = target_language.name_en[:2].lower()
17        source_lang_code = source_lang.name_en[:2].lower()
18
19        translated_plain_subs = translate(orig_plain_subs, "google",
20                                         target_lang_code, source_lang_code, "\n")
21        translated_subs = restore_webvtt_format(orig_subs, orig_plain_subs,
22                                                translated_plain_subs)
23
24    return Subtitles(
25        media_id=orig_base.media_id,
26        source_id=orig_base.source_id,
27        name=name, subtitles=translated_subs,
28        language=target_language,
29        season_number=orig_base.season_number,
30        episode_number=orig_base.episode_number)
31    return None

```

Рисунок Ж.3 – Лістинг функції перекладу

```

1 def get_synonyms(word: str) -> list[str]:
2     """ Get synonyms for a word using nltk. """
3     synonyms = set()
4     for syn in wordnet.synsets(word):
5         for lemma in syn.lemmas():
6             synonym = lemma.name().replace("_", " ")
7             if synonym.lower() != word.lower():
8                 synonyms.add(synonym)
9     return sorted(synonyms)

```

Рисунок Ж.4 – Лістинг функції отримання синонімів

```

1 @app.route(
2     "/api/available_subtitles/<source_id>/<id_on_source>",
3     methods=["GET"])
4 def get_available_subtitles(source_id: int, id_on_source: str):
5     """Returns list of available subtitles"""
6     media_manager = MediaManager()
7     media_id = media_manager.get_media_id(source_id, id_on_source)
8     if not media_id: return jsonify({"error": "Media not found"}), 404
9
10    subtitle_models = media_manager.get_media_subtitles(
11        media_id=media_id, include_text=True)
12
13    subtitles: List[dict] = [
14        model.model_dump(mode="python") for model in subtitle_models]
15    return jsonify(subtitles)

```

Рисунок Ж.5 – Лістинг функції отримання доступних субтитрів

```

1 @app.route(
2     "/api/get_categories/",
3     methods=["GET"])
4 def get_categories():
5     """Returns list of available subtitles"""
6
7     media_manager = MediaManager()
8     categories_models = media_manager.get_all_categories()
9
10    categories = [model.model_dump(mode="python")
11                  for model in categories_models]
12    return jsonify({"categories": categories})

```

Рисунок Ж.6 – Лістинг функції отримання створених словників

```

1 @app.route(
2     "/api/save_vocabulary",
3     methods=["POST"])
4 def save_vocabulary():
5     """Save vocabulary"""
6
7     data = request.json
8     original = data.get("original")
9     translated = data.get("translated")
10    examples = data.get("examples")
11    categories = data.get("categories")
12
13    examples_model = [Example(**y) for e in examples]
14    categories_model = [Category(**c) for c in categories]
15
16    vocabulary = Vocabulary(
17        original=original,
18        translated=translated,
19        examples=examples_model,
20        categories=categories_model)
21
22    media_manager = MediaManager()
23    media_manager.set_vocabulary(vocabulary=vocabulary)
24    return jsonify({"success": True})

```

Рисунок Ж.7 – Лістинг функції збереження лексичної одиниці

```

1 @app.route(
2     "/api/update_progress",
3     methods=["POST"])
4 def update_progress():
5     """Update progress of media watching"""
6
7     data = request.json
8     id_on_source = data.get("id_on_source")
9     source_id = data.get("source_id")
10    progress = data.get("progress")
11
12    media_manager = MediaManager()
13    media_id = media_manager.get_media_id(
14        source_id=source_id,
15        id_on_source=id_on_source)
16    media_manager.update_media_progress(
17        media_id=media_id,
18        progress=progress)
19    return jsonify({"success": True})

```

Рисунок Ж.8 – Лістинг функції фіксації поточної часової мітки

```

1 @app.route(
2     "/api/translate",
3     methods=["POST"])
4 def translate_text():
5     """Translate text"""
6
7     data = request.json
8     with_synonyms = data.get("synonyms")
9     text = data.get("text")
10    target_lang = data.get("target_lang", "uk")
11    source_lang = data.get("source_lang", "en")
12
13    if not text: return jsonify(
14        {"error": "No text provided"}), 400
15
16    translated = translate(
17        text=text,
18        translator_type="google",
19        target_language=target_lang,
20        source_language=source_lang)
21
22    synonyms = get_synonyms(word=text) if with_synonyms else None
23
24    return jsonify({"translated": translated,
25                  "synonyms": synonyms})

```

Рисунок Ж.9 – Лістинг функції отримання перекладу та синонімів

```

1 async function fetchAvailableSubtitlesApi() {
2     try {
3         const response = await fetch(
4             `${API_BASE_URL}/available_subtitles/1/${getMediaID()}`);
5         if (!response.ok)
6             throw new Error(
7                 `HTTP error! status: ${response.status}`);
8         availableSubtitles = await response.json();
9     } catch (error) {
10        availableSubtitles = null;
11        throw error;
12    }}

```

Рисунок Ж.10 – Лістинг функції отримання субтитрів із серверу

```

1 function createSubtitlesContent(subs_id) {
2     const subtitleText = createSubtitleContainer();
3     subtitleText.id = "sublearn-subtitle-text-" + subs_id;
4     subtitleText.classList.add("empty");
5
6     const contextMenu = createContextMenu(subs_id, subtitleText);
7     contextMenu.id = "sublearn-context-menu-" + subs_id;
8
9     subtitleContent = document.createElement("div");
10    subtitleContent.className = "subtitle-helper-content-container";
11
12    subtitleContent.appendChild(contextMenu);
13    subtitleContent.appendChild(subtitleText);
14
15    subtitleContainer.appendChild(subtitleContent);
16    setupTextSelectionHandler(subtitleText, contextMenu);
17 }

```

Рисунок Ж.11 – Лістинг функції, що вбудовує
контейнер субтитрів у DOM плеєра

```

1 async function updateProgressAPI(
2   id_on_source, source_id, progress
3 ){
4   try {
5     const response = await fetch(
6       `${API_BASE_URL}/update_progress`, {
7         method: "POST",
8         headers: { "Content-Type": "application/json" },
9         body: JSON.stringify({
10           id_on_source, source_id, progress
11         })
12     });
13
14     if (!response.ok)
15       throw new Error("Response was not ok");
16     return await response.json();
17   } catch (error) return null;
18 }

```

Рисунок Ж.12 – Лістинг функції оновлення
прогресу перегляду на сервері

```

1 async function getCategoriesAPI() {
2   try {
3     const response = await fetch(
4       `${API_BASE_URL}/get_categories/`, {
5         method: "GET",
6         headers: { "Content-Type": "application/json" }
7     });
8
9     if (!response.ok)
10       throw new Error("Response was not ok");
11
12     const data = await response.json();
13     return data.categories || [];
14   } catch (return []);
15 }

```

Рисунок Ж.13 – Лістинг функції отримання наявних словників

```

1 async function saveVocabularyAPI({
2   original, translated, examples, categories
3 }) {
4   try {
5     const response = await fetch(
6       `${API_BASE_URL}/save_vocabulary`, {
7         method: "POST",
8         headers: { "Content-Type": "application/json" },
9         body: JSON.stringify({
10           original, translated, examples, categories
11         })
12     });
13
14     if (!response.ok)
15       throw new Error("Response was not ok");
16     return await response.json();
17   } catch (error) return null;
18 }

```

Рисунок Ж.14 – Лістинг функції додавання одиниці лексики

```

1 async function translateTextAPI({
2   text, source_lang, target_lang, synonyms
3 }) {
4   try {
5     const response = await fetch(`${API_BASE_URL}/translate`,
6     {
7       method: "POST",
8       headers: { "Content-Type": "application/json" },
9       body: JSON.stringify({
10         text, source_lang, target_lang, synonyms
11       }),
12     });
13
14     if (!response.ok)
15       throw new Error("Response was not ok");
16     return await response.json(); // { translated, synonyms }
17   } catch return null;
18 }

```

Рисунок Ж.15 – Лістинг функції перекладу
та отримання синонімів

```

1 function createStatusIndicator() {
2   if (statusIndicator) statusIndicator.remove();
3   statusIndicator = document.createElement("div");
4   statusIndicator.className = "subtitle-helper-indicator";
5
6   statusIndicator.style.backgroundColor = (
7     subtitlesEnabled ? "#4CAF50" : "#F44336";
8   )
9   document.body.appendChild(statusIndicator);
10 }

```

Рисунок Ж.16 – Лістинг функції створення візуального
індикатору статусу активності розширення

```

1 def generate_test_questions(
2   media_type: str, media_title: str,
3   base_text: str, model: str = "gpt-4") -> str:
4   """Generate test questions based on the base text"""
5
6   prompt = f"""
7 generate test questions with answers. no extra words. format:
8 q: question.
9 (a-d) options.
10 a: answer.
11 {media_type}: {media_title}
12 base text:
13 {base_text}
14 """
15
16   response = openai.ChatCompletion.create(
17     model=model, messages=[
18       {"role": "user",
19        "content": prompt}
20     ],
21     temperature=0.7,
22     max_tokens=1000)
23
24   return response["choices"][0]["message"]["content"]

```

Рисунок Ж.17 – Лістинг функції створення візуального

Додаток И
Копія опублікованих результатів



ЗБІРНИК ТЕЗ ДОПОВІДЕЙ

Студентської науково-практичної конференції
ІНТЕЛЕКТУАЛЬНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В ПРИКЛАДНИХ
ДОСЛІДЖЕННЯХ
(ІТІАТ-2025)

27-29 травня 2025 року

Тернопіль

2025

Карнидал Володимир, Биковий Павло	155
ПРОГРАМНИЙ МОДУЛЬ ВИЗНАЧЕННЯ ВІКУ ТА СТАТІ ЛЮДИНИ НА ЗОБРАЖЕННІ ІЗ ЗАСТОСУВАННЯМ OPENCV ТА ГЛИБОКИХ НЕЙРОННИХ МЕРЕЖ	155
Касьян Тамара, Турченко Ірина	160
ПРОГРАМНИЙ МОДУЛЬ АНАЛІЗУ ТА ПРОГНОЗУВАННЯ ІНФЕКЦІЙНИХ ЗАХВОРЮВАНЬ ЗА ДОПОМОГОЮ МЕТОДІВ МАШИННОГО НАВЧАННЯ	160
Кацидим Віта, Ліп'яніна-Гончаренко Христина, Ралік Ігор	164
ІНТЕЛЕКТУАЛЬНИЙ АНАЛІЗ ЗАДОВОЛЕНOSTІ КЛІЄНТІВ ОБСЛУГОВУВАННЯМ НА ОСНОВІ МЕТОДІВ МАШИННОГО НАВЧАННЯ	164
Климчук Анастасія, Турченко Ірина	168
ІНТЕРАКТИВНИЙ ЗАСТОСУНОК ДЛЯ ВИВЧЕННЯ АНГЛІЙСЬКОЇ МОВИ НА ОСНОВІ МЕДІА-КОНТЕНТУ ІЗ ВИКОРИСТАННЯМ OPENAI	168
Ковальковський Віталій, Ліп'яніна-Гончаренко Христина	171
МОДУЛЬ РОЗПІЗНАВАННЯ ОБЛИЧЧЯ НА ОСНОВІ ГЛИБОКОГО НАВЧАННЯ	171
Ковтуненко Андрій, Ліп'яніна-Гончаренко Христина	176
TELEGRAM-БОТ ДЛЯ ВІДСТЕЖЕННЯ КУРСУ КРИПТОВАЛЮТ	176
Котов Владислав, Майків Ігор	179
МОДУЛЬ КЛАСИФІКАЦІЇ АНОТАЦІЙ НАУКОВИХ СТАТЕЙ З arXiv ЗА ДОПОМОГОЮ МОДЕЛІ RoBERTa	179
Кулик Степан	182
ВИЯВЛЕННЯ МЕРЕЖЕВИХ АТАК У ТРАФІКУ ВЕЛИКИХ ДАНИХ НА ОСНОВІ ЕВОЛЮЦІЙНИХ ОБЧИСЛЕНЬ	182
Мельничук Руслана, Ліп'яніна-Гончаренко Христина	185
МОДУЛЬ ОПТИМІЗАЦІЇ МАРКЕТИНГОВИХ КАМПАНІЙ ЗА ДОПОМОГОЮ КЛАСТЕРИЗАЦІЇ КЛІЄНТІВ	185
Матвійшин Наталія, Майків Ігор	188
МОДУЛЬ АНАЛІЗУ ТА КЛАСИФІКАЦІЇ СЮЖЕТІВ ФІЛЬМІВ ЗА ЖАНРАМИ З ВИКОРИСТАННЯМ БІБЛІОТЕКИ SPACY	188
Мукан Павло, Лендюк Тарас	191
МОДУЛЬ ВИЯВЛЕННЯ МОВИ ВОРОЖНЕЧІ З ВИКОРИСТАННЯМ БІБЛІОТЕКИ SCIKIT-LEARN ДЛЯ КЛАСИФІКАЦІЇ ТЕКСТУ	191
Наконечна Ірина, Ліп'яніна-Гончаренко Христина	194
ВИЗНАЧЕННЯ ДЖЕРЕЛА ТЕКСТУ ЗА ДОПОМОГОЮ НЕЙРОМЕРЕЖІ НА ОСНОВІ EMBEDDING І LSTM	194

ІНТЕРАКТИВНИЙ ЗАСТОСУНОК ДЛЯ ВИВЧЕННЯ АНГЛІЙСЬКОЇ МОВИ НА ОСНОВІ МЕДІА-КОНТЕНТУ ІЗ ВИКОРИСТАННЯМ OPENAI

У сучасних умовах розвитку цифрових технологій знання іноземних мов є необхідним для доступу до медіа з оригінальних джерел без мовних бар'єрів. Дефіцит часу та кліпове мислення зумовлюють потребу в автоматизованих й ефективних методах самоосвіти. Особливо актуальним є використання медіа-контенту, що поєднує навчання з відпочинком і забезпечує глибше занурення в мовне середовище [1].

Медіа-контент як освітній інструмент персоналізує навчання і природно вписується у повсякденне життя. Аудіювання дозволяє чути живу мову носіїв у реальних ситуаціях, що зменшує акцент і розвиває такі мовні навички як артикуляцію, інтонацію і ритм. Регулярна практика в автентичному контексті допомагає сприймати мову на рівні її внутрішньої структури та логіки [2].

Враховуючи зазначені вище аргументи та переваги і недоліки проаналізованих відомих аналогів «Cake» [3] і «Listen in English» [4], у межах дослідження створено застосунок, що забезпечує інтерактивний процес навчання та систематизацію мовного матеріалу. Функціонал доповнюється браузерним розширенням для роботи із субтитрами. Згідно алгоритму (рисунок 1) застосунок пропонує чотири основні напрямки його використання: взаємодію з медіа-контентом, опрацювання лексики, аналітику та управління резервними даними.

Розроблений застосунок складається з модулів: «Media Management», «Exercises», «Vocabulary», «Analytics», «Extension» та «User Component». Основним є модуль роботи з медіа, що надає доступ до навчального відео, втілює функціонал керування й навігації по ньому та інтегрується із зовнішніми сервісами перегляду. Для обробки даних із сайтів з захистом від ботів застосовано бібліотеку `curl_cffi.requests` [5]. Підтримку кількох джерел субтитрів реалізовано через шаблон «Фабричний метод», що забезпечує гнучкість та модульність архітектури [6]. Також тут реалізовано функцію генерації перекладу субтитрів з наявних текстів за допомогою Google і MyMemory сервісів з очищенням від службових символів. Переклад підтримує паралельне виконання потоків, кількість яких адаптується до ресурсів системи.

Модуль опрацювання лексики відповідає за керування лексичними даними та особистими словниковими наборами. Також він спрямований на розвиток навичок письма й аудіювання. Перше реалізується через інтервальні картки, друге – через адаптивні аудіо-тести, що генеруються ШІ на основі переглянутого контенту. Для стабільності результатів запити було оптимізовано: уточнено інструкції, зменшено кількість токенів, забезпечено чіткий формат «запитання-відповідь» із урахуванням типу та назви відео.

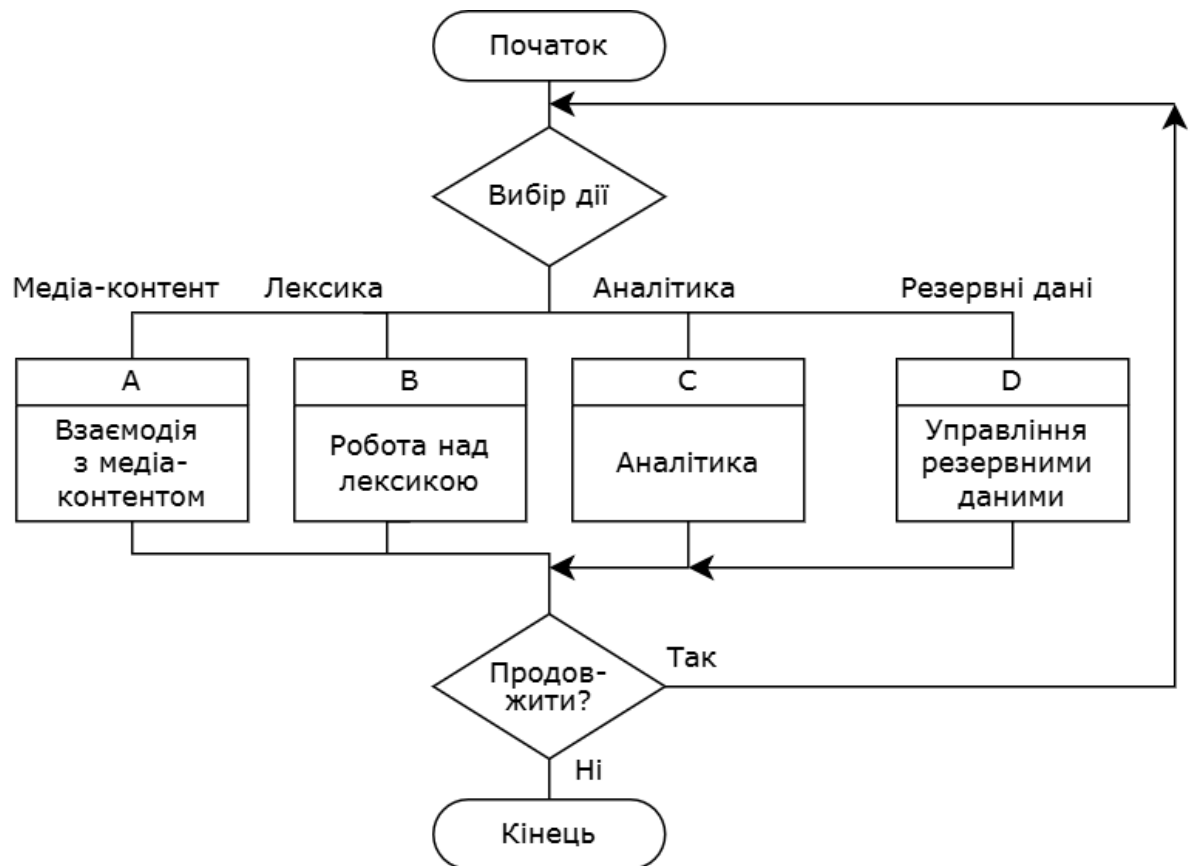


Рисунок 1 – Загальна схема алгоритму роботи інтерактивного застосунку

Модуль аналітики відслідковує результати тестування і засвоєння лексики, формуючи показники прогресу. На основі типів завдань обчислюються ключові метрики ефективності: активність, кількість правильних відповідей, повторень.

Модуль резервного копіювання забезпечує збереження та відновлення історії прогресу.

Браузерне розширення доповнює функціональність основного застосунку під час перегляду відео на зовнішніх платформах. Воно дозволяє: перемикати режими субтитрів; перекладати обрані фрагменти; додавати лексику до словника.

Інтерфейс користувача взаємодіє із сервером та розширенням для роботи із субтитрами. Графічна оболонка охоплює всі ключові функціональні блоки системи. Міжмодульна взаємодія забезпечує обмін контентом, результатами тестів і статистикою.

У рамках дослідження реалізовано багатофункціональний застосунок для вивчення лексики в контексті, що: поєднує перегляд цільового відео-контенту, інтерактивну роботу із субтитрами, навчальні вправи та інструменти ІІІ. Він сприяє самоосвіті, стимулюючи формування стійкої мотивації до навчання. Перспективними напрямками розвитку є підтримка мобільних платформ, розширення типів вправ, робота з локальними відео, прикріплення власних субтитрів та інтеграція із ширшим спектром стримінгових сервісів, включно без вбудованих субтитрів.

Список використаних джерел

1. Porter S. I., Castillo M. S. The effectiveness of immersive language learning: An investigation into English language acquisition in immersion environments versus traditional classroom settings. *Research Studies in English Language Teaching and Learning*. 2023. Vol.1(3), P.155–165.
2. Liu Y., Zhang X., Zhang J. Investigation of Facilitating English Speaking of Elementary School Students in Authentic Contexts with UEnglish. *Human-Centric Intelligent Systems*. 2023. Vol.3(1), P.1–11.
3. Cake. URL: <https://cake.day/>.
4. Listening. URL: <https://learnenglish.britishcouncil.org/skills/listening>.
5. curl_cffi. URL: <https://pypi.org/project/curl-cffi/>.
6. Factory Method. URL: <https://refactoring.guru/design-patterns/factory-method>.