

Міністерство освіти і науки України
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»
Кафедра мультимедійних систем факультету інформатики

Розробка веб-застосунку для вивчення іноземних мов

**Текстова частина до курсової роботи
за спеціальністю «Інженерія програмного забезпечення» 121**

Керівник курсової роботи
Старший викладач
Борозенний С. О.

(підпис)
“ ” 2022 р.

Виконав студент 4-го курсу
Кузан О.О.
“ ” 2022 р.

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ «КИЄВО-МОГИЛЯНСЬКА АКАДЕМІЯ»

Кафедра мультимедійних систем факультету інформатики

ЗАТВЕРДЖУЮ

Зав. кафедри мультимедійних систем,

доцент, к.ф.-м.н.

О.П. Жежерун
(підпис)

“ ____ ” _____ 2021 р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

на курсову роботу

студенту _____ Кузану Олегу _____

____ 4-го _____ курсу факультету інформатики

ТЕМА: Розробка веб-застосунку для вивчення іноземних мов

Вихідні дані:

- Веб-додаток для тренування іншомовної лексики

Зміст ТЧ до курсової роботи:

Вступ

1. Аналіз предметної області. Постановка завдання курсової роботи
2. Теоретичні відомості
3. Опис реалізації програмного продукту

Висновки

Список джерел

Додатки (за необхідністю)

Дата видачі “ ____ ” _____ 2021 р.

Керівник _____ Завдання отримано _____

Календарний план виконання курсової роботи

Тема: Розробка клієнт-серверної системи діагностування пацієнтів

Календарний план виконання роботи:

№ п/п	Назва етапу курсової роботи	Термін виконання етапу	Примітка
1.	Отримання завдання на курсову роботу	Жовтень- листопад 2021р.	
2.	Вивчення та аналіз задачі	Листопад- грудень 2021р.	
3.	Розробка архітектури та загальної структури програми	Січень- лютий 2022р.	
4.	Створення застосунку	Лютий- березень 2022р.	
5.	Написання текстової частини	Березень-квітень 2022р.	
6.	Перегляд курсової роботи науковим керівником	Травень 2022р.	
7.	Захист курсової роботи		

Студент Кузан О. О. _____

Керівник Борозенний С. О. _____

“ _____ ” _____ 2021 р.

Зміст

Список іншомовних термінів	6
Анотація.....	7
Вступ.....	8
1 Аналіз предметної області	9
1.1 Огляд ринку flash-card застосунків.....	9
1.1.1 Reword	9
1.1.2 Reverso Context	10
1.1.3 Toucan.....	12
1.1.4 ABBYY Lingvo.....	12
1.1.5 Quizlet	13
2.1 Аналіз ринку.....	14
2 Розробка концепту застосунку	15
2.1 Формулювання принципів.....	15
2.1.1. Гейміфікація.....	15
2.1.2 Контроль.....	15
2.1.3. Доступність	16
2.1.4. Пріоритизація.....	16
2.1.5. Взаємодія.....	16
2.1.6. Дизайн	17
2.1.7. Лінгвістичний контекст	17
2.2 Підготовка до реалізації	18
2.2.1 Datamuse	18
2.2.2 Stanford CoreNLP	18

2.2.3 Oxford	19
2.2.4 Wordnik.....	19
2.2.5 Words API.....	19
2.2.6 WordNet.....	20
2.2.7 Urban dictionary	20
2.2.8 Owlbot.....	20
2.2.9 Twinword	20
2.2.10 Інше	21
3 Розробка додатку.....	22
3.1 Обґрунтування вибору засобів розробки.....	22
3.2. Опис реалізації	22
3.2.1 Логін та реєстрація	22
3.2.2 Discover	23
3.2.2 Профіль	24
3.2.3 Тренування	25
3.2.4 Налаштування	26
Висновок.....	27
Список використаних джерел.....	28

Список іншомовних термінів

Flashcard	- флеш-картки для запам'ятовування, на кшталт стікерів
POS-tagging	- визначення частини мови
Front-end	- видима у браузері сторона застосунку
MWE	- термін, що складається з кількох слів
Back-end	- серверна сторона застосунку
Endpoint	- адреса сервісу, на яку можна робити запити
Gamification	- використання ігрових практик у неігровому середовищі
Syllabification	- правила переносу слів на письмі
Голонім	- Y – голонім X, якщо X – частина Y
Меронім	- Зворотній до голоніма
Гіперонім	- Y – гіперонім X, якщо X – різновид Y
Гіпонім	- зворотній до гіпероніма
Word family	- сім'я спільнокореневих слів
Гомофони	- слова, що звучать однаково
Лематизація	- приведення словоформи до початкової форми
Іммерсивний	- з «зануренням» у середовище
API	- прикладний програмний інтерфейс

Анотація

Розглянуто існуючі рішення в сфері додатків для розширення та тренування лексикону, проаналізовано їх сильні сторони та недоліки, сформульовано концепт нового продукту, який би був більш універсальним, ніж уже існуючі.

Ключові слова: MWE, flash-card, web-app, Java, Angular, immersive learning.

Вступ

Вивчення іноземних мов – було, є, і буде актуальним, і це точно не зміниться «anytime soon», адже не зважаючи навіть на різноманітні глобальні потрясіння, свій об'єднаний, як ніколи до цього в історії – в тому числі, і навіть в першу чергу, мова йде про інформаційний вимір. Ми живемо в одному інфополі з людьми, які фізично можуть знаходитися за тисячі кілометрів від нас, але тим не менше, якщо ми розуміємо один одного, ця відстань не має значення – ми можемо вчитись і працювати разом, радіти на перемогою улюблених команд, насолоджуватись тими самим художніми творами. Все це можливо завдяки інтернету, але самої цифрової близькості недостатньо – кооперація стає на порядок продуктивнішою, якщо ви можете спілкуватись однією мовою.

Метою цієї курсової роботи є концептуальна та фактична розробка веб-застосунку для асистування у розширенні та закріпленні словникового запасу просунутого користувача, зроблена на основі аналізу функціоналу наявних рішень у сфері додатків для самостійно вивчення іноземних мов.

Робота складається з трьох розділів. У першому розглянуто кілька яскравих лідерів ринку, описано головний функціонал, проаналізовано сильні та слабкі сторони. У другому – на базі попереднього дослідження та особистого досвіду використання такого роду засобів, сформульовано концепт застосунку, який би як і міг поєднати в собі ключові риси уже існуючих рішень, так і надати потужніше машинно-лінгвістичне забезпечення процесу, опираючись на бази даних і програмні інтерфейси відомих провайдерів. У третьому розділі описана фактична реалізація концепту, описаний весь цикл існування лексики для користувача нашого сервісу – від поля вводу до успішного запам'ятовування, а також супутній функціонал.

1 Аналіз предметної області

1.1 Огляд ринку flash-card застосунків

1.1.1 Reword

Mobile-застосунок доступний для користувачів Андроїд та iOS. Простий та надійний інструмент. У процесі розвитку додатку розширювались заготовані списки слів, покращувалась озвучка, для полегшення запам'ятовування почали пропонувати асоціативні картинки, додався автопереклад. Доступний для англійської та ще восьми мов. Слова можна завантажити у додаток з файлу у форматі «слово; транскрипція (опціонально); переклад». Так, можна частково компенсувати відсутність клієнта «Reword» для десктопу, але очевидно це рішення далеко не для всіх.

```

511 concoct ; kənk'akt ; выдумывать, изобретать, сочинять, сострять
512 concoction ; kənk'akʃən ; приготовление
513 concubine ; k'ənkjəb,aɪn ; наложница
514 concur ; kənk'ə ; сходиться в мнениях, приходит к общему выводу, соглашаться
515 concurrently ; kənk'zəntʃli ; одночасно
516 concussion ; kənk'ʌʃən ; удар, сотрясение, взбалтывание
517 condemn ; kənd'ɛm ; осуждать, порицать
518 condolence ; kənd'ʊʃəns ; соболезнование, сочувствие
519 condone ; kənd'ʊn ; мириться, попустительствовать, потворствовать
520 conducive to ; kənd'ʊsɪv t'u ; благоприятный
521 confer on ; ; придают, возлагают на

```

Рисунок 1.1 – Файл для імпортування в Reword

Слово-картка має 6 ітерацій, часові проміжки між ними складають 30 хвилин, дві години, один день, три дні, 14 днів, 60 днів. На 3+ спробі при неправильній відповіді ітератор скидається на 3 спробу. Відповідь можна надати одним з трьох способів: надрукувати, вибрати із чотирьох запропонованих варіантів, чи просто відкрити відповідь. Чи пускати картку на наступну ітерацію користувач вирішує сам. Є мінімальна гейміфікація – лічильник днів, у які користувач займався підряд, а також позначається, чи виконана ціль на день.

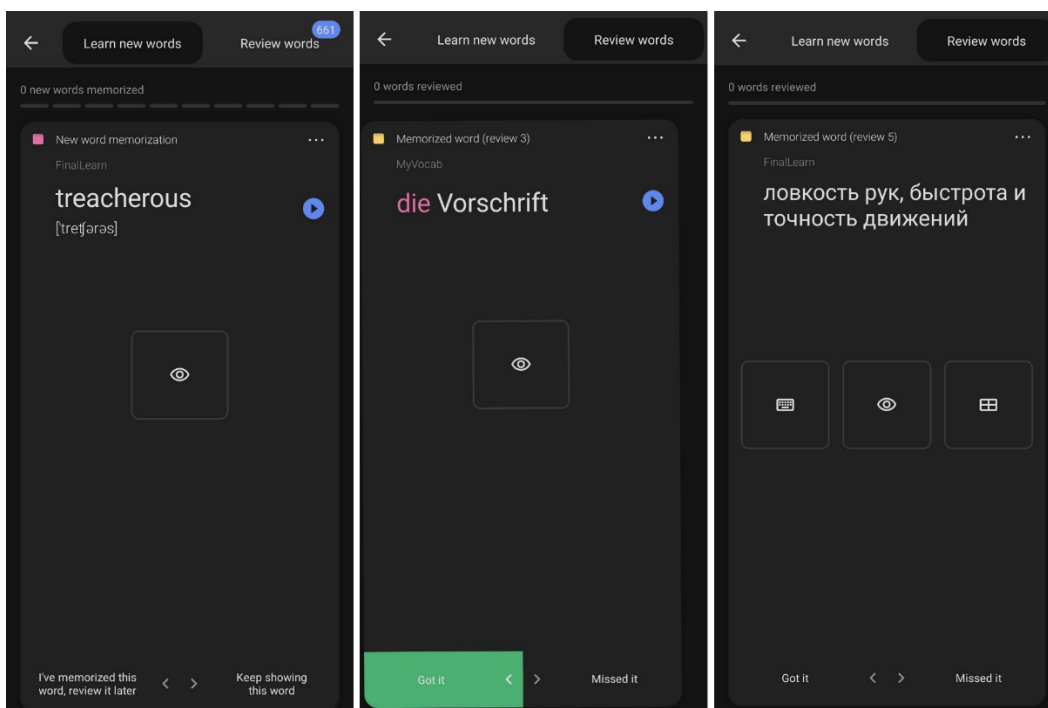


Рисунок 1.2 – Інтерфейс додатку Reword

1.1.2 Reverso Context

«Reverso» підтримує більше 10-и мов і є незамінним інструментом при навчанні, адже дає для шуканих слів та фраз реальний контекст із фільмів, серіалів, книг, а також художній переклад цих уривків. Це дозволяє зрозуміти, як в реальному житті використовується даний конструкт, і як його перекладали справжні перекладачі, дає відчуття відтінків значень та самому розгледіти певні паттерни, що не можливо зробити, маючи на руках лише сухі уривки з академічних словників. Ось приклад виводу:

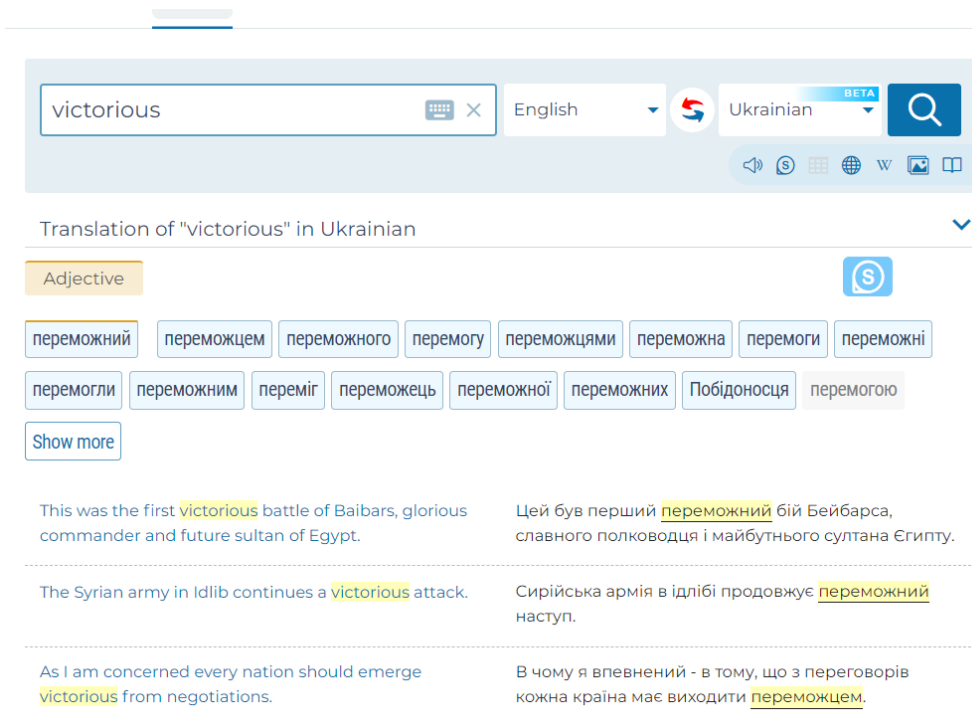


Рисунок 1.3 – Інтерфейс веб-сайту Reverso

Ще однією унікальною спроможністю «Reverso» є нещодавно доданий переклад субтитрів «on-the-fly», ось як це виглядає, наприклад, в YouTube:

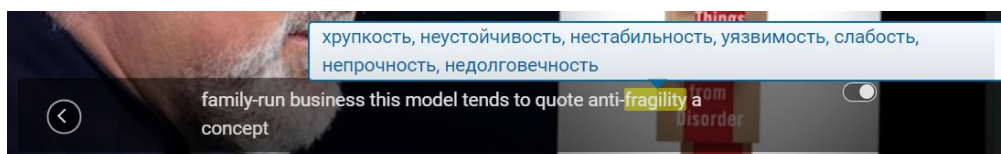


Рисунок 1.4 – Переклад субтитрів у Reverso

При наведенні курсора на ненативну плашку відео застигає, кожне слово можна перекласти у спливаючому вікні «Reverso». Така доступність в один клік не просто приємний досвід користувача, а необхідна умова для іммерсивного навчання, коли користувач занурюється в іншомовне середовище, і, наприклад, вивчає і практикує англійську розважаючись чи насолоджуючись улюбленим серіалом.

1.1.3 Toucan

Є ще один додаток, ідея якого теж пов'язана з таким «фоновим» навчанням. «Toucan» реалізований у вигляді розширення для браузера. Під час звичайного серфінгу в мережі, деякі слова прямо в тексті будуть замінені на відповідні переклади в цільовій мові, таким чином додаток ненав'язливо вкраплює лексику, яку користувач підхопить несвідомо і без цілеспрямованих зусиль. Такі втручання можна «приспати» на певний час в налаштуваннях, або скоригувати інтенсивність вкраплень.

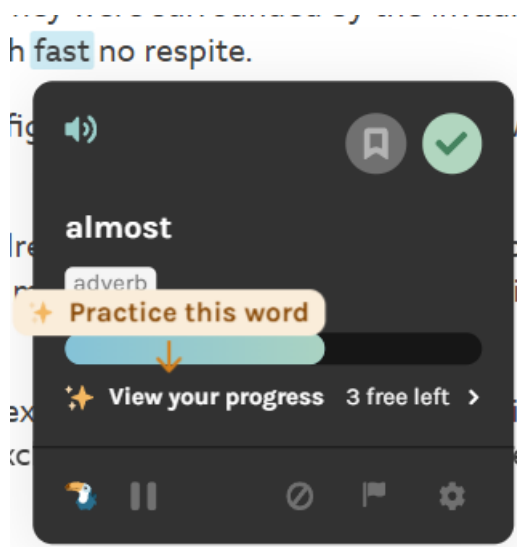


Рисунок 1.5 – Toucan в дії

1.1.4 ABBYY Lingvo

Найпотужнішу лінгвістичну підтримку мабуть надають засоби ABBYY Lingvo, в якому вбудовані багато відомих словників. Проект оперує рекордною кількістю мов (20) та словників (220), зокрема *Collins Cobuild Advanced Learner's English Dictionary*. Багато тематичних та вузькоспеціалізованих словників, які стосуються, наприклад, гірничодобувної справи. Проект доступний для смартфонів під брендом *Language Live*, як офлайнове десктопне рішення ABBYY Lingvo x6, є також веб-версія Lingvo Live. Слабкою стороною є тренування –

переклади можна відзначити "зірочкою" і додати в улюблені, але простим подальшим переглядом цих карток функціонал і обмежився.

1.1.5 Quizlet

Має режим флеш-карток, ідентичний Reword'у. Теж є заготовлені колекції та озвучка до них. З відмінностей:

- Є можливість повернутися назад до попередньої картки;
- Крім того, є режим «Match» - де треба поєднати слово з тлумаченням
- Є режим тесту, де способи підтвердження (істина/хиба, тестовий вибір і тд.) чергуються. В кінці можна оглянути введені відповіді, проаналізувати помилки;
- Є можливість створити персоналізований план - в залежності наскільки добре треба вивчити матеріал та наскільки ви з ним уже ознайомлені.
- Нема ніяких функціональностей специфічних до вивчення мови.

За підсумками можемо скласти таку таблицю:

Таблиця 1.1 – Порівняння додатків для вивчення мов

Сервіс	On-the-fly	Плат-форми	Алгоритм	Availability	Ціна	Мови
ABBY Lingvo	+/-	Bci	Список	Freemium	1120€	20
Quizlet	-	Bci	Базовий	Freemium	3\$ / міс	-
Reword	-	Mobile	Базовий	Freemium	100€	9
Lingvist	-	Bci	Просунутий	Premium	3\$ / міс	13
Reverso	+	Bci	Список	Freemium	94€ / міс	14
Toucan	+	Браузер	Просунутий	Freemium	2.99\$ / міс	11

2.1 Аналіз ринку

Як бачимо, в відкритому доступі знаходяться дуже багато різних за ідеями та виконанням додатків, які фокусуються на тих чи інших аспектах опанування іноземною мовою. Але ті, які дають мають сильнішу лінгвістичну базу, зазвичай не забезпечують якісний процес тренування і запам'ятовування своїх матеріалів, а ті, які сфокусовані на флоу флеш-карток, не мають достатньо потрібного лінгвістичного функціоналу. Також у додатках такого типу не реалізований потенціал взаємодії користувачів, яка би служила хорошим мотиватором для занять, підвищила би шанс не закинути навчання.

2 Розробка концепту застосунку

2.1 Формулювання принципів

2.1.1. Гейміфікація

Гейміфікація на перший погляд може здатись дитячою та зайвою, але цей аспект уже здобув належне визнання, і його роль стала предметом багатьох наукових досліджень, всі з яких сходяться на тому, що це ефективний засіб, який підвищує засвоюваність інформації та мотивує учасників. Цим простим трюком не потрібно нехтувати, бо найкраще навчання – це гра.

2.1.2 Контроль

Я переконаний, що користувачеві потрібно надавати максимальний контроль над процесом, адже ніхто не може знати краще, що для нього працює, ніж він сам. Це стосується в першу чергу кінцевого вибору «згадав» / «не згадав» на карточці, який треба завжди залишати за користувачем, причиною хибно-негативної відповіді може бути помилкове натискання чи помилка вводу, а прикладом хибно-позитивного результату може бути збиття порядку карточок, або перемикання режимів, коли гравець мав змогу бачити загадуване слово за кілька хвилин до перевірки, що звичайно не вкладається в прийнятні часові рамки для одної ітерації. Одним словом, користувач завжди правий. Додаток не має бути екзаменатором, а лише знаряддям в руках студента.

Абсолютні значення з будь-якого аспекту процесу за можливості потрібно виносити на розсуд користувачеві, як чисельні (кількість ітерацій), так і логічні (як-от вибір лінгвістичних сервісів), не потрібно вгадувати за користувача те, про що можна запитати у нього самого. В той же час, ми можемо і нам слід вказувати рекомендовані наперед задані значення, аби не перевантажувати юзерів, особливо новачків. Таким чином, матимемо адаптивну кастомізацію – яка підійде і користувачам, яким використовують сервіс в стилі "take and go", або тим, яким кортить все підлаштувати під себе.

2.1.3. Доступність

Оскільки наш продукт задумується як знаряддя-асистент для іммерсивного навчання, критично необхідно, щоб процес від знаходження нового слова в його природному середовищі до оформлення картки нашому додатку можна було пройти максимально швидко, безболісно і за мінімальну кількість кліків. Різниця в кілька секунд може бути межею, яка лежить за зоною комфорту користувача, якому потрібно відірватися від інфо-продукту для збереження якоїсь лексичної одиниці. У застосунка має бути браузерне розширення для швидкого захоплення слова/виразу і введення його у систему.

2.1.4. Пріоритизація

Потрібно давати користувачу при навчанні більшу картину, ніж просто слово-визначення. Тут ключовим параметром є частота зустрічання лексичного входження та визначення рівня користувача при реєстрації на спеціальному тесті. Відповідно до нього користувач має приймати проінформоване рішення, чи актуальне для нього це відносно його знань. При повторенні першими мають йти входження, які з більшою ймовірністю зустрінуться в реальному тексті, та ті, які більш релевантні в контексті рівня користувача. В іншому випадку зусилля користувача будуть витрачатись на слова, які він не зможе зустріти в посильних йому матеріалах в реальному житті, що результує в демотивацію та розчарування.

2.1.5. Взаємодія

Змога користувача ділитись своїми досягненнями та прогресом з друзями буде позитивним моментом не тільки для нього/неї, але й для самої платформи - людський компонент (наприклад «скільки користувачів разом з вами вчать це слово») спонукатиме проводити на ній більше часу, надасть відчуття пов'язаності з реальним світом. Елемент «змагання» також є позитивним фактором, який

можна розвинути в рамках гейміфікації. Навчатись завжди веселіше разом. Але неізолюваність навчальних процесів може бути корисною і з технічної точки зору – накопичивши достатню кількість даних, можна використовувати їх для фасилітації навчання наступних користувачів. Зокрема може йтися про часті помилки в написанні, які допускали попередні студенти, збір метрики, дані про те, наскільки відносно складно давався цей вираз тощо. Можна також створити внутрішній банк готових карток чи цілих колекцій, які користуються популярністю. Можливість ділитись з друзями картками виведе флеш-картковий застосунок на новий рівень.

2.1.6. Дизайн

Застосунок повинен бути візуально приємним для користувача, не перевантажений непотрібними деталями, з консистентними застосування тем і кольорових гам, мати впізнаваний стиль. Не лише функціонал має приваблювати користувачів, а й сам вигляд і досвід використання. Так процес навчання буде викликати менше внутрішнього спротиву, який неодмінно буває, адже навчання це зазвичай рутинна праця і вихід з зони комфорту. Flashcard застосунок з лексиною це доволі концентрована суміш знань, яку важко перетравити самотужки.

2.1.7. Лінгвістичний контекст

Додана вартість додатку не в останню чергу має полягати в тому, аби опираючись на API та бібліотеки відомих словників, енциклопедій, сервісів, і баз даних, було надано всі необхідні метадані, які стосуються введеної користувачем інформації. Це також зможе направити новачка, якому важко виокремити інфінітивну невідмінену базову форму із речення самому, або який сплутав частини мови, стикнувся з нерегулярною множиною тощо. Вхідження запам'ятовуються тільки в контексті, а не в вакуумі чи самі по собі, тому має бути

надана широка вибірка "живих" використань. Також підвищить продуктивність виведення при наявності цілого гнізда слів (word family) - вивчивши базове слово, юзер впорається і з спільнокореневими теж. Потрібно заохочувати запам'ятовувати слово в якомога ширшому контексті, на кшталт "це слово є частина ось такої популярної фрази. Не хочете вивчити її натомість?".

2.2 Підготовка до реалізації

Для надання потрібного контексту до лексичних входжень необхідно послуговуватись якимись спрямованими на це сервісами. В рамках курсової роботи було розглянути багато таких пропозицій на ринку, як безкоштовних так і платних. Їхній функціонал задає рамку можливостей нашого застосунку.

2.2.1 Datamuse

АПІ для розробників, яке також базується на WordNet та Wictionary, сфокусовано на різноманітні фільтри значення, звучання, написання тощо. Можемо знайти слова, які звучать як наше, які часто передують або слідують за нашим, синоніми, антоніми, частотність, гомофони, частину мови. Рахує, але не виводить кількість складів, у полі метадані можна вивести інформацію про визначення (під капотом відбувається лематизація, дані з WordNet), POStagging, вимова, та частотність (дані з Google Books Ngrams).

2.2.2 Stanford CoreNLP

Набір інструментів під Java для обробки природних мов (Natural Language Processing), дає нескінченні можливості для кастомізації дослідження мов для ентузіастів. Наразі підтримує 8 мов, а саме: арабська, китайська, англійська, французька, німецька, угорська, італійська та іспанська. У повного CoreNLP API є укорочена спрощена версія Simple API. Виконує точний POS-tagging, за яким розрізняє 3 позицій, а також лематизацію.

2.2.3 Oxford

Один з найбільш поціновуваних словників англійської мови, стандарт якості. Пропонує лематизацію, переклади на багато мов, приклади, фрази з даним словом, частину мови, вимову з аудіофайлом, працює зі складеними термами.

2.2.4 Wordnik

Wordnik API може похизуватися визначеннями з 5-ох словників: American Heritage Dictionary of the English Language, the Century Dictionary, Wiktionary, the GNU International Dictionary of English, і з Wordnet'у, які покривають понад 800,000 англійських слів. Також дає синоніми, антоніми, приклади, аудіо, частотність, етимологію, вимову, пов'язані фрази та випадкове слово дня. Ключ видається протягом тижня всім бажаючим. За допомогою флагу "useCanonical" можна діставати вивід по леммі слова, а не якісь його форми.

2.2.5 Words API

Словник та тезаурус. Переклади та взаємозв'язки бере з WordNet, а частотність аналізувалась по даних з Open Subtitles. Має ендпойнти also - повертає phrasal verbs від даного дієслова, та similarTo - слова, схожі за значенням, але не синоніми. Є також ендпойнт на частотність, повертає відносне значення від 1 до 7, чим більше число, тим частіше зустрічається. Надає ендпойнти для знаходження голонімів, меронімів, гіпернімів, гіпонімів, але в реальному житті лише малий відсоток слів буде мати такого типу реляції. Ендпойнт на вимову дає слова у ролі кожної частини мови, у якій він зустрічається.

2.2.6 WordNet

«WordNet», який, як ми побачили, часто використовується як фундамент інших API, розроблявся при Пристонському університеті з 1985 року. Це поєднання словника та тезауруса. Слова у ньому розбиті на synsets - групи когнітивних (повних) синонімів, які пов'язані між собою лексичними та семантичними зв'язками. Ця БД вільно розповсюджується і доступна для завантаження.

2.2.7 Urban dictionary

Urban dictionary - словник-збірник фраз англomовного сленгу, неологізмів, тобто того, чого зазвичай не знайдеш в стандартизованих словниках. Тим не менше, словник дає хороші приклади і тлумачення і до звичайних слів.

2.2.8 Owlbot

Owlbot - маленький проект, надає інформацію взятую з індексування відкритих джерел. Нема обмежень, токен дається за запитом на пошту, але і видає доволі скупі результати з єдиного ендпойнту: вимова, приклади, визначення, частина мови.

2.2.9 Twinword

Фокусується на аналізі цілих текстів, серед його API наприклад, вловлення емоційного підтексту, схожості змістів тощо. Надає два корисні в нашому контексті ендпойнти – перший це лематизація, інший це виведення відносної складності слова по шкалі від одного до семи. Оскільки сервіс один з найдорожчих з розглянутих, ці дані були співставленні з більш доступним «Datamuse», який надає абсолютні числа, для переведення їх у відносну шкалу такого типу, а лематизація була делегована «Stanford CoreNLP».

2.2.10 Інше

У «Reverso Context», на жаль, не має публічного API, тому його не інтегруємо, а просто відсилаємо запит користувача на сторонній ресурс, який відривається в новій вкладці. Так і не вдалось отримати доступ до «Collins Dictionary». Для розшифрування акронімів використовується Acromine REST API від NaCTeM.

Можемо так узагальнити зібрану інформацію про мовні API:

. Таблиця 2.1 – Порівняння мовних API

API	Free quota	Лемма тизація	Частотні сть	MWE	Силабифікація	Визначення, синоніми	Ціни
Oxford	1K / міс.	+	-	+	-	+	£0.0022
Free Dictionary	Необм.	-	-	+	-	+	-
Urban	Необм.	-	-	+	-	+	-
Datamuse	100K / д.	-	+	+	+/-	+	-
Words API	2500 / д.	-	+	+	+	+	10\$/ міс
Wordnik	100 / год.	+/-	+	-	+	+	10\$/ міс
Owlbot	Необм.	-	-	-	-	+	-
Twinword	500 / міс	+	+	+	-	-	\$0.003

3 Розробка додатку

3.1 Обґрунтування вибору засобів розробки

Back-end частина була написана на Java 11, з використанням фреймворку Spring Boot, для front-end'у був обраний фреймворк Angular (11-ої версії). Ця комбінація добре зарекомендувала себе в індустрії та широко використовується у найрізноманітніших застосунках. Проект збирався за допомогою Maven, за контроль версій відповідав Git, база даних – PostgreSQL на AWS RDS.

3.2. Опис реалізації

3.2.1 Логін та реєстрація

Логін та реєстрація відбуваються в одному вікні. Для реєстрації необхідно надати ім'я користувача (унікальне), пошту (унікальна) та пароль. Є опція зареєструватись або залогінитись через Google провайдера використовуючи технологію OAuth2.0. Це все робить процес автентифікації максимально комфортним.

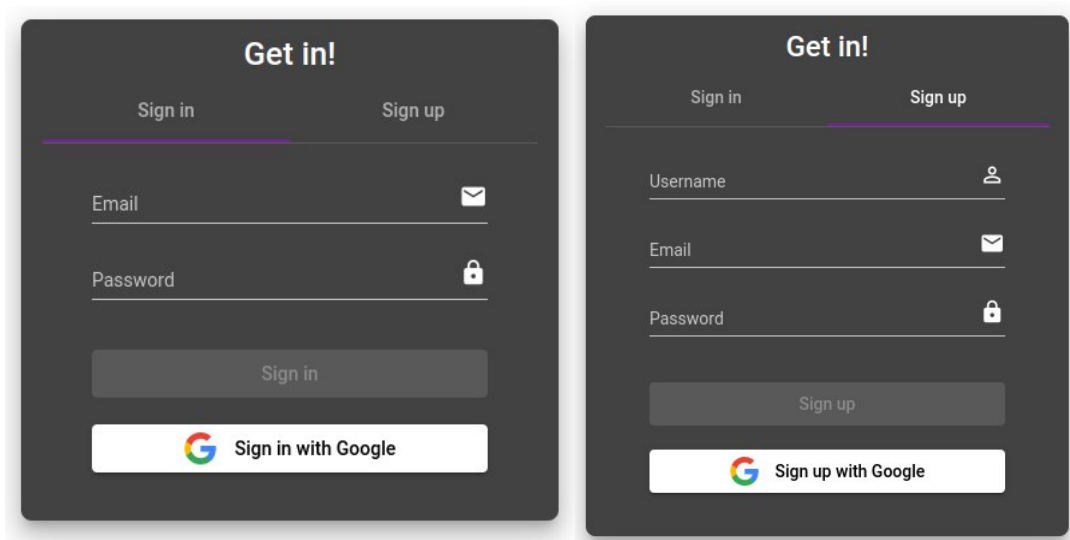


Рисунок 3.1 – Логін та реєстрація

3.2.2 Discovery

На головному вікні програми - секції «Discover», є велике поле вводу з автодоповненням, варіанти до якого вантажаться із наймісткішого списку всіх англійських слів, які вдалось знайти у мережі, він налічує 462 тисячі одиниць. Доповнення іде до кожного слова, довшого ніж 4 символи після останнього пробілу, фрази не пропонуються.

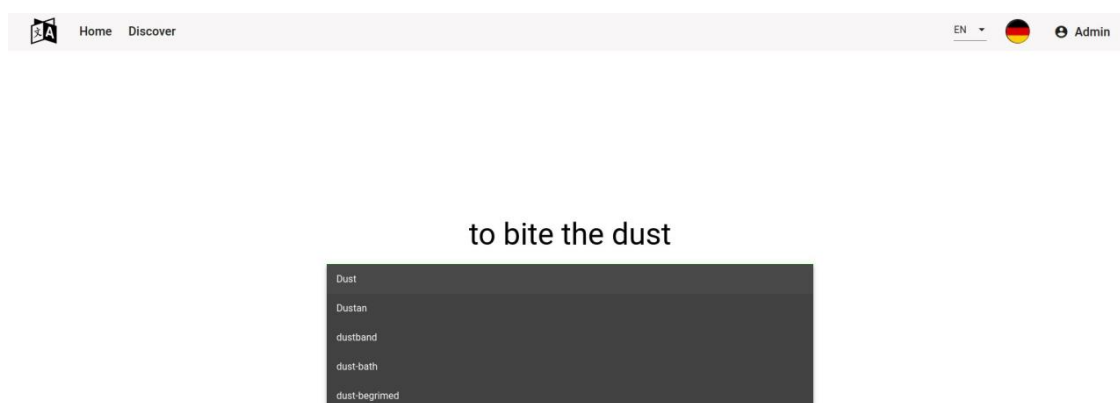


Рисунок 3.2 – Вікно Discover з автодоповненням

Після натискання «Enter» дані відправляються на наш сервер, і починається лінгвістична обробка - лематизація, силабіфікація, тощо. Виводиться швидкий результат з прикладами, вимовою, та можливістю пошукати в окремо обраних двигунах – «Reverso», «Urban», «Oxford». Якщо слово (або слово з родинного гнізда) було знайдено в колекціях користувача, йому виводять знайдене входження, де він може додати новий приклад. Тобто «Discovery» використовується не тільки для додавання нових слів, але і для швидкого пошуку по власних колекціях, можливо з метою внесення змін, нових прикладів.

Користувач має можливість створити картку. Її можна заповнити перетягнувши приклади з пошукових двигунів, вказавши рівень зацікавленості/пріоритетності для себе (за промовчанням NORMAL, може бути змінений на LOW або ж HIGH). В нотатках можна занотувати якусь інформацію на майбутнє.

Для ведення статистики можна заповнити поле «source» – з якого джерела ця словоформа («YouTube», «Netflix», «News», щось інше). Також можна вказати, чи ви хочете це слово мати в активному лексиконі (тобто могли згадати при нагоді), чи пасивному (могли зрозуміти при наступні зустрічі). В залежності від цього перемикача слово буде в тренуванні зараховуватись при введенні або оригіналу (перший варіант), або перекладу (другий). Користувачу, який в налаштуваннях вказав, скільки часу він проводить в іноземному інформаційному середовищі, виводяться дані, як часто при такому режимі йому зустрінеться це слово.

Лінгвістичний аналіз може підказати користувачу те, що він насправді шукав чи мав на увазі, наприклад, базову форму слова. Можна клацнути на пропозицію (лему слова, наприклад) і цим запустити процес заново.

3.2.2 Профіль

На сторінці профіль можна переглянути і редагувати власні дані, як-от пошта, пароль, ім'я користувача. Тут також є секція друзів: їх можна знайти за імейлом.

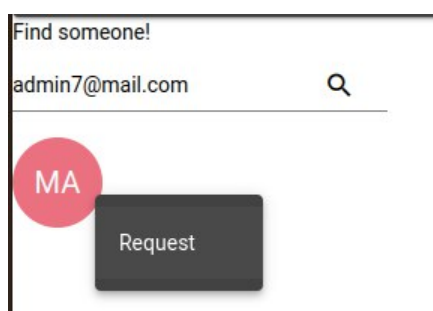


Рисунок 3.3 – Пошук друзів

Після того, як ви надішлете запрошення, ваша іконка з'явиться у друга у вкладці «Incoming requests», а у вас відповідно у вкладці «Outcoming requests», Цю заявку можна підтвердити (тоді ви будете друзями), скасувати (запит видалиться з БД), а також заблокувати користувача (його не будуть сповіщати,

але більше надокучати запитами він не зможе). Ви можете бачити, кого заблокували, але не навпаки. З друзями можна ділитись своїми карточками, колекціями, порівнювати рекорди і статистику.

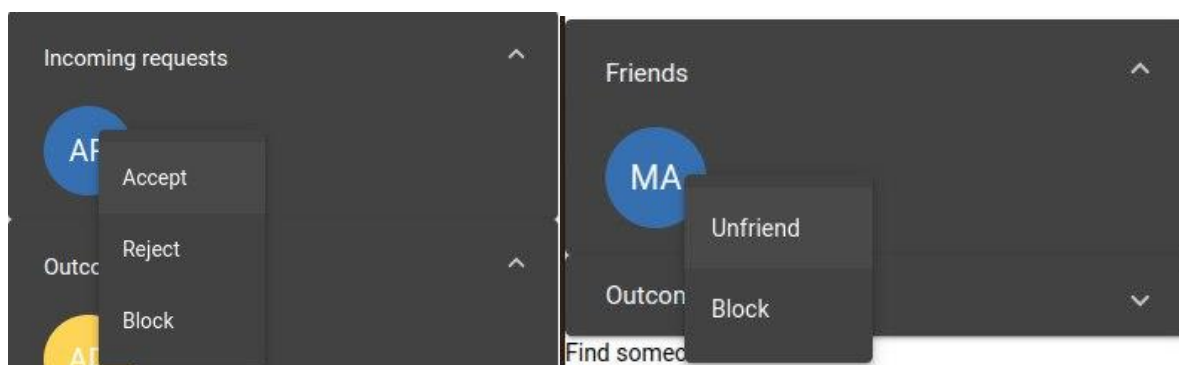


Рисунок 3.4 – Контексте меню для друзів та заявок

3.2.3 Тренування

На вкладці тренування свій стек карток можна сортувати за запропонованим алгоритмом (поєднує кілька метрик), за власною пріоритетністю (HIGH, NORMAL, LOW), яка задається при створенні (за промовчанням NORMAL), за частотністю, за датою створення. Користувачеві виводяться дані: скільки у нього слів на кожній ітерації, загальна статистика вивчених. Можна повторювати також за номером ітерації – наприклад, перше розібратись з тими словами, які готовий іти на фінальний огляд.

На етапі перегляду при неправильному написанні, система запропонує вам надалі підкреслювати проблемне місце, збіг приголосних чи голосних, наприклад. В залежності від сприйняття вами входження (entry), можна відредагувати очікуваний бар'єр проходження - підвищити його, якщо відчуваєте, що слово дається важко. При перегортанні слів діє «завмирання» (freeze) на кілька секунд, аби навіть легке слово не було спокуси перегорнути зашвидко, так воно закарбується у пам'яті. Зручним інструментом є теги: чому у вас не виходить згадати? Можливо потрібно більше прикладів, більше контексту. Додайте

мнемонічну фразу для кращого запам'ятовування в нотатках. Позначте написання/звучання як нетипове – ці дані будуть використані, щоб попередити про це інших користувачів в майбутньому. Самі обирайте спосіб підтвердження - тестовий вибір, друкування в полі вводу, чи відкривати картку без перевірок, кінцеве рішення приймати вам. Якщо замість одного слова ви ввели інше, яке також вивчаєте, система запропонує вам позначити ці слова як «плутані» (confused), щоб надалі ви звертали на них особливу увагу.

3.2.4 Налаштування

На вкладці «Settings» можна відкоригувати багато різних параметрів:

- тривалість «завмирання» на тренуваннях;
- кількість ітерацій за промовчанням;
- не приймати нових запитів в друзі;
- провайдер прикладів за промовчанням;
- вимкнути силабіфікацію;
- вимкнути вивід лінгвістичного контексту;
- вимкнути підсвітку подвоєнь літер;
- вимкнути звуковий супровід;
- вимкнути гейміфікацію;
- показувати першим слово / переклад / як вказано на картці;
- включити режим «Strict spelling».

«Strict spelling» означає, що введенне має абсолютно точно співпадати з входженням. Під завмиранням мається на увазі мінімальний час, який картку має бути на екрані до того, як її перегорнуть.

Висновок

В рамках роботи було протестовано функціонал популярних сучасних застосунків на ринку вивчення іноземних мов, сформульовано концепт проекту, у якому були би заповнені функціональні пробіли наявних реалізацій та запропоновано новий підхід до розширення лексикону, базуючись на власному багаторічному досвіді використання наявних рішень та дослідженню лінгвістичної складової предметної області.

Можна окреслити такі напрямки продовження дослідження та розробки: додати підтримку для інших популярних мов (локалізації та власне навчання), розширити можливість внутрішньо-платформової взаємодії користувачів, покращити лінгвістичний аналіз та систему рекомендацій, розробити браузерне розширення для швидкого зберігання нових слів, а також відточувати UI/UX аспект.

Список використаних джерел

1. Анатолій Загнітко, «Словник Сучасної Лінгвістики: Поняття І Терміни», ДонНУ, Донецьк, 2012
[Електронний ресурс]
<https://r.donnu.edu.ua/bitstream/123456789/248/1/%D1%81%D0%BB%D0%BE%D0%B2%D0%BD%D0%B8%D0%BA%20%D1%81%D1%83%D1%87%D0%B0%D1%81%D0%BD%D0%BE%D1%97%20%D0%BB%D1%96%D0%BD%D0%B3%D0%B2%D1%96%D1%81%D1%82%D0%B8%D0%BA%D0%B8%20-%201%20%D0%9E%D0%A1%D0%A2%2B%2B.pdf>
2. Davletbaeva D.N., «Lectures on English lexicology», Kazan university, 2010
[Електронний ресурс]
<https://kpfu.ru/docs/F1797492221/Lectures.on.Lexicology1.pdf>
3. Sofia Rebelo, Pedro Isaías, «Gamification as an Engagement Tool in E-Learning Websites», Journal of Information Technology Education. Research; Santa Rosa, 2020 [Електронний ресурс]
<https://www.proquest.com/openview/043078067a7ccde3b3ce289a478e2350/1?pq-origsite=gscholar&cbl=5324456>
4. Yutaro Shigeto, Ai Azuma, Sorami Hisamoto, Shuhei Kondo, Tomoya Kose, Keisuke Sakaguchi, Akifumi Yoshimoto, Frances Yung, Yuji Matsumoto, «Construction of English MWE Dictionary and its Application to POS Tagging», Nara Institute Science of Technology (NAIST), 2013 [Електронний ресурс]
<https://aclanthology.org/W13-1021.pdf>
5. Akihiko Kato, Hiroyuki Shindo, and Yuji Matsumoto, «Construction of an English Dependency Corpus incorporating Compound Function Words. In Proceedings of the Tenth International Conference on Language Resources and Evaluation», European Language Resources Association (ELRA), 2016. [Електронний ресурс]
<https://aclanthology.org/L16-1263.pdf>