

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»

Факультет систем управління літальних апаратів

Кафедра комп'ютерних наук та інформаційних технологій

Пояснювальна записка
до кваліфікаційної роботи
(тип кваліфікаційної роботи)

магістра

(освітній ступінь)

на тему «Розробка web-застосунку для підготовки до сертифікації на знання іноземної
мови»

ХАІ.302.367.23О.122.193096 ПЗ

Виконав: студент 6 курсу групи № 367
Спеціальність 122 «Комп'ютерні науки»

(код та найменування)

Освітня програма Комп'ютеризація обробки
інформації та управління

Кочев Д.Д.

(прізвище й ініціали студента)

Керівник: Прохоров О.В.

(прізвище й ініціали)

Рецензент: Метешкін К.О.

(прізвище й ініціали)

Харків – 2023

Міністерство освіти і науки України
Національний аерокосмічний університет ім. М.Є. Жуковського
«Харківський авіаційний інститут»

Факультет систем управління літальних апаратів
Кафедра комп'ютерних наук та інформаційних технологій
Рівень вищої освіти другий (магістерський)
Спеціальність 122 «Комп'ютерні науки»
Освітня програма Комп'ютеризація обробки інформації та управління

ЗАТВЕРДЖУЮ
Завідувач кафедри
Олег ФЕДОРОВИЧ

(підпис) (ім'я та прізвище)
«07» листопада 2023 р.

З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ СТУДЕНТА

Кочеву Денису Дмитровичу
(прізвище, ім'я, по батькові)

1. Тема кваліфікаційної роботи «Розробка web-застосунку для підготовки до сертифікації на знання іноземної мови»

керівник кваліфікаційної роботи Прохоров О.В., д.т.н., проф., проф. каф. 302,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом Університету № 1968-уч. від «06» листопада 2023 року

2. Термін подання студентом кваліфікаційної роботи 10.01.2024.

3. Вихідні дані до проекту (роботи) хмарна документо-орієнтовна база даних MongoDB, хмарне об'єктне сховище Google Cloud Storage, API великої мовної моделі ChatGPT 3.5-turbo, API системи розпізнавання тексту Whisper, програмне забезпечення для пошуку помилок з відкритим кодом Language Tool, API моделей озвучення тексту ElevenLabs.

4. Зміст пояснювальної записки (перелік завдань, які потрібно розв'язати) аналіз особливостей та проблем мовної сертифікації, проектування web-застосунку для підготовки до мовної сертифікації, розробка алгоритмів оцінки письмових та усних завдань за допомогою методів штучного інтелекту для обробки природної мови, програмна реалізація web-застосунку, опис режимів роботи web-застосунку.

5. Перелік графічного матеріалу Розробка web-застосунку для підготовки до сертифікації на знання іноземної мови; Ціль та задачі дослідження; Аналіз потреби у мовних компетентностях; Огляд ринку застосунків підготовки до мовної сертифікації; Огляд існуючих рішень; Архітектура існуючих рішень; Головні недоліки існуючих рішень; Технічні вимоги до застосунку; Використовувані засоби та технології; Структура системи; Автоматизація оцінки письмових та усних завдань; Процес розробки клієнтської частини; Режим роботи застосунку; Development and Operations; Результати дослідження; Дякую за увагу.

6. Консультанти розділів кваліфікаційної роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
з основної частини	Прохоров О.В. д.т.н., проф. каф. 302		
зі спеціального розділу	Прохоров О.В. д.т.н., проф. каф. 302		
з економіки та управління	Малєєва Ю.А., доц. каф. 302		

Нормоконтроль _____ Олександр ПРОХОРОВ «____» _____ 20__ р.
(підпис) (ім'я та прізвище)

7. Дата видачі завдання « 31 » жовтня 20 23 р.

КАЛЕНДАРНИЙ ПЛАН

№ п/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів кваліфікаційної роботи	Примітка
1.	Аналіз проблем мовної сертифікації в контексті автоматизації підготовки до іспитів	31.10.23-05.11.23	
2.	Концептуальне проектування модулів застосунку для підготовки до мовної сертифікації.	06.11.23-12.11.23	
3.	Автоматизація обчислень результатів виконання мовних сертифікаційних іспитів.	13.11.23-21.11.23	
4.	Програмна реалізація функціональних модулів системи.	22.11.23-15.12.23	
5.	Економічний розділ.	25.12.23-27.12.23	
6.	Тестування та розгортання системи.	27.12.23-04.01.24	
7.	Оформлення пояснювальної записки.	31.10.23-09.01.24	

Студент _____ Денис КОЧЕВ
(підпис) (ім'я та прізвище)

Керівник кваліфікаційної роботи _____ Олександр ПРОХОРОВ
(підпис) (ім'я та прізвище)

РЕФЕРАТ

Розробка web-застосунку для підготовки до сертифікації на знання іноземної мови. – Кочев Денис Дмитрович, кваліфікаційна робота магістра, Харків, Національний аерокосмічний університет ім. М.Є. Жуковського «ХАІ», кількість сторінок 114, кількість малюнків 63, кількість таблиць 19, кількість джерел літератури 15.

КЛЮЧОВІ СЛОВА: штучний інтелект, обробка природньої мови, хмарні обчислення, токенізація текстових даних, контейнеризація програмних модулів.

ЦІЛЬ: автоматизація процесу підготовки до сертифікаційних іспитів на знання іноземної мови.

ОБ’ЄКТ ДОСЛІДЖЕННЯ: процеси автоматизації обробки природньої мови у форматі тексту або аудіо.

ПРЕДМЕТ ДОСЛІДЖЕННЯ: методи, моделі та засоби штучного інтелекту для обробки природньої мови для автоматизації підготовки до сертифікаційних іспитів на знання іноземної мови.

ЗАДАЧІ: аналіз особливостей та проблем мовної сертифікації; проектування web-застосунку для підготовки до мовної сертифікації; розробка алгоритмів оцінки письмових та усних завдань з використанням методів штучного інтелекту для обробки природньої мови; програмна реалізація web-застосунку; впровадження безперервної інтеграції, розгортання та опис режимів роботи web-застосунку.

МЕТОДИ ДОСЛІДЖЕННЯ: методи проектування баз даних, методи штучного інтелекту.

ОТРИМАНІ РЕЗУЛЬТАТИ: розроблено web-застосунок для підготовки до сертифікації на знання іноземної мови з використанням методів штучного інтелекту для обробки природньої мови.

ABSTRACT

Development of the web-application to prepare for foreign language certification. – Kochev Denys, Master's Thesis, Kharkiv, National Aerospace University "Kharkiv Aviation Institute", amount of pages 114, amount of figures 63, amount of tables 19, amount of sources of literature 15.

KEY WORDS: artificial intelligence, natural language processing, cloud computing, text tokenization, application containerization.

OBJECTIVE: automatization of preparation processes for language certification exam.

OBJECT OF RESEARCH: procedures of natural language processing automatization.

SUBJECT OF RESEARCH: artificial intelligence methods, models and tools for natural language processing to automate preparation for foreign language certification exams.

TASKS OF RESEARCH: analysis of the features and issues of language certification; system design of web application for language certification preparation; development of evaluation algorithms for writing and speaking tasks with artificial intelligence; programming implementation of the web application; deployment, testing and review of the product.

METHODS OF RESEARCH: database design techniques, artificial intelligence methods for software engineering.

RESULTS OF RESEARCH: web application for foreign language certification preparation with artificial intelligence is fully implemented and available.

ЗМІСТ

ВСТУП	8
1 АНАЛІЗ ПРОБЛЕМ МОВНОЇ СЕРТИФІКАЦІЇ В КОНТЕКСТІ АВТОМАТИЗАЦІЇ ПІДГОТОВКИ ДО ІСПИТІВ.....	9
1.1 Сучасні технології та мовні сертифікаційні іспити.....	9
1.2 Вивчення потреби в сервісах для підготовки до мовної сертифікації	11
1.3 Аналіз технічного стану існуючих рішень	15
1.3.1 Огляд платформи Lingua	16
1.3.2 Огляд інтернет-ресурсу Lerntipp2.....	18
1.3.3 Аналіз платформи Sprachtest.....	19
1.3.4 Огляд веб-застосунку Sprachportal	21
1.4 Висновки щодо технічного стану наявних рішень	23
2 КОНЦЕПТУАЛЬНЕ ПРОЕКТУВАННЯ МОДУЛІВ ЗАСТОСУНКУ ДЛЯ ПІДГОТОВКИ ДО МОВНОЇ СЕРТИФІКАЦІЇ.....	26
2.1 Особливості проектування в контексті предметної області.....	26
2.2 Проектування сховища	28
2.3 Проектування головного модуля.....	35
2.4 Проектування функціональних модулів	37
2.5 Висновки щодо проектування системи.....	42
3 ОСОБЛИВОСТІ АВТОМАТИЗАЦІЇ В ОБЧИСЛЕННЯХ РЕЗУЛЬТАТІВ МОВНИХ СЕРТИФІКАЦІЙНИХ ІСПИТІВ	44
3.1 Алгоритмізація оцінки письмових завдань	44
3.2 Інтеграція моделей штучного інтелекту в систему	48
3.3 Автоматизація усної частини сертифікаційного іспиту.....	51
3.4 Висновки до розділу з автоматизації	54
4 ПРОГРАМНА РЕАЛІЗАЦІЯ МОДУЛІВ СИСТЕМИ ДЛЯ ПІДГОТОВКИ ДО СЕРТИФІКАЦІЇ НА ЗНАННЯ ІНОЗМЕННОЇ МОВИ	56
4.1 Вибір інструментальних засобів	56
4.2 Розробка головного модуля.....	60
4.3 Програмування модуля обробки та аналізу тексту.....	63
4.4 Висновки за розділом.....	71
5 ВПРОВАДЖЕННЯ КОНВЕЙЕРУ БЕЗПЕРЕРВНОЇ ІНТЕГРАЦІЇ ТА ОГЛЯД ФУНКЦІОНАЛЬНИХ МОЖЛИВОСТЕЙ ЗАСТОСУНКА.....	73
5.1 Впровадження автоматичного тестування	73

5.2 Огляд головних функціональних можливостей системи	77
5.3 Розгортання застосунку	85
5.4 Висновки за розділом	87
6 ЕКОНОМІЧНА ЧАСТИНА	88
6.1 Мета розділу	88
6.2 Визначення тривалості та розрахунок собівартості виконання проекту (перший варіант)	88
6.3 Визначення тривалості та розрахунок собівартості виконання проекту (другий варіант)	95
6.4 Висновки за розділом	99
ВИСНОВКИ	100
ПЕРЕЛІК ПОСИЛАНЬ	101
ДОДАТОК А. Графічні матеріали	103

ВСТУП

Актуальність теми дослідження. Високий інтерес до платформ для підготовки до мовних іспитів та зростаюча потреба ІТ-спеціалістів в покращенні своїх мовних навичок корелює з глобалізацією ринку праці та збільшенням мобільності спеціалістів. Дані тенденції свідчать про актуальність і потребу в розробці ефективних сервісів. Використовуючи сучасні моделі штучного інтелекту з'являється можливість автоматизувати ряд перевірок, яких частково або повністю позбавленні існуючі сервіси для підготовки до мовних сертифікаційних іспитів.

Метою дослідження є автоматизація процесу підготовки до сертифікаційних іспитів на знання іноземної мови.

Основні завдання дослідження: аналіз особливостей та проблем мовної сертифікації; проектування web-застосунку для підготовки до мовної сертифікації; розробка алгоритмів оцінки письмових та усних завдань з використанням методів штучного інтелекту для обробки природньої мови; програмна реалізація web-застосунку; впровадження безперервної інтеграції, розгортання та опис режимів роботи web-застосунку.

Об'єктом дослідження є процеси автоматизації обробки природньої мови у форматі тексту або аудіо.

Предметом дослідження є методи, моделі та засоби штучного інтелекту для обробки природньої мови для автоматизації підготовки до сертифікаційних іспитів на знання іноземної мови.

Практичне значення отриманих результатів проведеного дослідження полягає в можливості використання розробленої платформи в процесі вивчення німецької мови або підготовки до сертифікаційних іспитів. На момент захисту кваліфікаційної роботи застосунок надає готові рішення для підготовки німецьких сертифікаційних іспитів рівня B1. Застосунок доступний за публічною адресою <https://einfache-sprachen.de/>.

Науково-технічний ефект полягає в автоматизації оцінки результатів виконання письмових та усних частин мовного сертифікаційного іспиту з використання моделей та методів штучного інтелекту для обробки природньої мови.

Структура і обсяг кваліфікаційної роботи магістра. Кваліфікаційна робота магістра складається з вступу, 6 розділів, висновків і 1 додатку. Повний обсяг кваліфікаційної роботи магістра становить 114 сторінок, в тому числі: 63 рисунка по тексту; 19 таблиць по тексту; списку використаної літератури з 15 найменувань на 2 сторінках; 1 додаток на 12 сторінках.