

ХЕРСОНСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ДИЗАЙНУ
КАФЕДРА ПРОГРАМНИХ ЗАСОБІВ І ТЕХНОЛОГІЙ

Пояснювальна записка

до кваліфікаційної роботи магістра

на тему:

**«ДОСЛІДЖЕННЯ МЕТОДІВ ПРОГРАМНОЇ РЕАЛІЗАЦІЇ
СИНХРОННОГО ГОЛОСОВОГО ПЕРЕКЛАДУ В РЕАЛЬНОМУ ЧАСІ»**

Виконав: здобувач 6 курсу, групи 6ПР
спеціальності 121 «Інженерія
програмного забезпечення»

Хміль Олег Олегович

(прізвище та ініціали)

Керівник: к.т.н., доцентка Огнєва О.Є.

(прізвище та ініціали)

Рецензент к.т.н., доц. Корніловська Н.В.

(прізвище та ініціали)

Хмельницький – 2025 р

Факультет Інформаційних технологій та дизайну
Кафедра Програмних засобів і технологій
Освітньо-кваліфікаційний рівень Магістр
Галузі знань 12 «Інформаційні технології»
Спеціальність 121 «Інженерія програмного забезпечення»
Освітньо-професійної програми «Програмна інженерія»

ЗАТВЕРДЖУЮ

Завідувач кафедри ПЗіТ,

професор О.Є.Огнева

«___» _____ 2025 року

З А В Д А Н Н Я

НА ДИПЛОМНУ РОБОТУ ЗДОБУВАЧУ

Хміль Олег Олегович

(прізвище, ім'я, по батькові)

1. Тема роботи «Дослідження методів програмної реалізації синхронного голосового перекладу в реальному часі»

керівник роботи Огнева Оксана Євгенівна к.т.н., доцентка,

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

затвержені наказом ХНТУ від «15» вересня 2025 року № 416-с.

2. Строк подання здобувачем роботи _____

3. Вихідні дані до роботи Матеріали навчальної-виробничої практики

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити)
Дослідження предметної області та постановка задачі, аналіз методів автоматичного розпізнавання мовлення, машинного перекладу та синтезу мовлення, проєктування програмного забезпечення, тестування програмного забезпечення

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)

1. Flowchart-діаграма

2. DFD-діаграма

3. UML-діаграма послідовності

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Нормоконтроль			
Антиплагіат			

7. Дата видачі завдання 01.10.2025

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного роботи	Строк виконання етапів роботи	Примітка
1	Збір матеріалу за темою дипломного проектування; дослідження предметної області, визначення та узгодження індивідуальних тем кваліфікаційних робіт	01.10.2025-05.10.2025	Готово
2	Теоретичне дослідження методів ASR, NMT, TTS, аналіз інформаційних технологій	06.10.2025-20.10.2025	Готово
3	Системний аналіз і обґрунтування проблеми	21.10.2025-31.10.2025	Готово
4	Методи та засоби проектування прототипу	01.11.2025-10.11.2025	Готово
5	Проектування та реалізація прототипу системи	11.11.2025-21.11.2025	Готово
6	Написання вступу, висновків, оформлення списку джерел та додатків. Оформлення пояснювальної записки	11.11.2025-21.11.2025	Готово
7	Перевірка на плагіат, нормоконтроль, отримання рецензій та інших супровідних документів.	01.12.2025-12.12.2025	
8	Підготовка до захисту та захист роботи	15.12.2025	

Здобувач _____ Хміль О.О. _____

(підпис)

(прізвище та ініціали)

Керівник роботи _____ Огнева О.Є _____

(підпис)

(прізвище та ініціали)

АНОТАЦІЯ

Тема кваліфікаційної роботи: «Дослідження методів програмної реалізації синхронного голосового перекладу в реальному часі».

Автор проєкту: Хміль Олег Олегович.

Керівник проєкту: Огнєва Оксана Євгеніївна.

Пояснювальна записка: 86 с., 8 рис., 40 джерел.

Графічна частина: 12 слайдів.

СИНХРОННИЙ ПЕРЕКЛАД, РОЗПІЗНАВАННЯ МОВЛЕННЯ, NLP

Метою кваліфікаційної роботи є дослідження та розробка програмної системи синхронного голосового перекладу в реальному часі, що забезпечує перетворення мовлення у текст, переклад та синтез зворотного голосового повідомлення. У процесі роботи було проведено аналіз сучасних методів обробки мовлення та асинхронного перекладу, визначено вимоги до системи та спроектовано програмний застосунок.

У роботі застосовано технології розпізнавання мовлення на основі нейронних мереж, методи машинного перекладу та синтезу мовлення. Використано сучасні програмні засоби та API, такі як Whisper ASR, Google Translate API та інші інструменти NLP.

Результатом роботи є рішення, що виконує синхронний переклад мовлення в реальному часі, має зручний інтерфейс, підтримує декілька мов та забезпечує стабільну якість перекладу.

Дата

Підпис

ANNOTATION

Thesis topic: "Research of Methods for Software Implementation of Real-Time Synchronous Voice Translation"

Author: Khmil Oleh Olehovych

Supervisor: Ohnieva Oksana Yevheniivna

Explanatory note: 86 pages, 8 figures, 40 sources

Graphical part: 12 slides

SYNCHRONOUS TRANSLATION, SPEECH RECOGNITION, NLP

The purpose of the qualification work is to research and develop a software system for real-time synchronous voice translation that provides speech-to-text conversion, translation, and reverse voice message synthesis. During the work, an analysis of modern speech processing methods and asynchronous translation was conducted, system requirements were defined, and a software application was designed.

The work applies neural network-based speech recognition technologies, machine translation methods, and speech synthesis. Modern software tools and APIs such as Whisper ASR, Google Translate API, and other NLP tools were used.

The result of the work is a solution that performs real-time synchronous speech translation, features a user-friendly interface, supports multiple languages, and provides stable translation quality.

Data

Signature

ЗМІСТ

ВСТУП.....	
РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ.....	
1.1 Сутність та принципи синхронного голосового перекладу.....	
1.2 Сучасні методи розпізнавання, перекладу та синтезу мовлення.....	
1.3 Аналіз технологій реального часу.....	
1.4 Висновки до розділу 1.....	
РОЗДІЛ 2. ПОШУК ГОТОВИХ РІШЕНЬ І ЇХ ПОРІВНЯННЯ.....	
2.1 Огляд існуючих систем синхронного перекладу.....	
2.2 Порівняльний аналіз технологій і архітектур.....	
2.3 Аналіз та вибір технологій і методів реалізації.....	
2.4 Висновки до розділу 2.....	
РОЗДІЛ 3. ПРОГРАМНА РЕАЛІЗАЦІЯ.....	
3.1 Архітектура системи синхронного перекладу.....	
3.2 Проектування інтерфейсу користувача.....	
3.3 Реалізація модулів.....	
3.4 Взаємодія модулів та обробка в реальному часі.....	
3.5 Висновки до розділу 3.....	
РОЗДІЛ 4. ТЕСТУВАННЯ ТА АНАЛІЗ РЕЗУЛЬТАТІВ.....	
4.1 Методика тестування функціональності застосунку.....	
4.2 Тестування швидкодії та якості перекладу.....	
4.3 Верифікація та валідація системи.....	
4.4 Висновки до розділу 4.....	
ВИСНОВКИ.....	
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	
ДОДАТКИ.....	

ВСТУП

У сучасну епоху глобалізації, стрімкого розвитку інформаційних технологій та розширення цифрового простору людей з різних країн об'єднує постійний обмін інформацією. Міжнародні комунікації стали невід'ємною частиною щоденного життя, а багатомовність перестала бути рідкісним явищем, що стало нормою. У цьому контексті питання подолання мовних бар'єрів набуває особливої актуальності, оскільки ефективне спілкування є ключовим чинником успіху у сфері науки, бізнесу, освіти, туризму, дипломатії та соціальних сервісів. Технології синхронного перекладу, що раніше існували лише у вигляді спеціалізованого обладнання або фантастичних концепцій, сьогодні стають доступними для широкого кола користувачів завдяки інтелектуальним алгоритмам та мобільним пристроям.

Швидкий розвиток штучного інтелекту, зокрема у галузі обробки природної мови, дозволив перейти від традиційних статистичних моделей перекладу до глибоких нейронних мереж, які здатні аналізувати контекст, розуміти структуру речень і забезпечувати надзвичайно точний і природний переклад у реальному часі. У цьому контексті однією з головних тенденцій стало поширення систем Neural Machine Translation, що поєднують можливості нейронних мереж із масштабованими хмарними сервісами. Поряд із цим значного розвитку зазнали технології Automatic Speech Recognition, що забезпечують точне перетворення голосового сигналу в текст, а також системи Text-to-Speech, які дають змогу синтезувати природне мовлення [12].

Технології, які ще десять років тому були доступні лише великим корпораціям, сьогодні інтегруються у смартфони, планшети, розумні колонки та інші пристрої. Моделі на зразок OpenAI Whisper демонструють високу точність розпізнавання мовлення різними мовами, навіть коли запис

здійснюється у шумних або нестандартних умовах. Google Translate API забезпечує швидку обробку тексту з можливістю перекладу на десятки мов, а Firebase дає змогу легко зберігати та синхронізувати користувацькі дані. Усі ці фактори створюють ідеальні передумови для розроблення мобільної системи синхронного голосового перекладу в реальному часі, яка є зручною, гнучкою, масштабованою та готовою до використання у різних життєвих ситуаціях, що можуть виникнути [12].

Актуальність теми зумовлена стрімким розвитком глобальних цифрових комунікацій, у межах яких ефективна взаємодія між людьми різних мовних спільнот стає ключовою умовою успішної роботи у багатьох сферах від бізнесу та освіти до медицини й міжнародної співпраці. У сучасному світі потреба у швидкому та точному подоланні мовного бар'єра зростає щодня, оскільки більшість інформації передається у реальному часі, а затримка в комунікації може стати критичною. Саме це робить системи синхронного голосового перекладу не просто корисним інструментом, а стратегічно важливою технологією.

Мета і завдання дослідження цієї роботи полягає в створенні реальної програмної системи, яка здатна здійснювати повний цикл обробки голосу: запис, розпізнавання, переклад, синтез мовлення і відтворення. Така система повинна працювати на мобільних пристроях у режимі реального часу та забезпечувати стабільний результат незалежно від умов використання. Для досягнення мети необхідно дослідити сучасні моделі розпізнавання мовлення, зокрема модель Whisper, проаналізувати можливості сервісів автоматичного перекладу, таких як Google Translate API, а також оцінити якість різних алгоритмів синтезу мовлення. Важливо також вивчити особливості Flutter, як кросплатформного середовища розробки, адже від правильного вибору архітектури залежить продуктивність, масштабованість і подальший розвиток цих системи.

Сучасні тенденції демонструють, що інтелектуальні системи перекладу поступово інтегруються в такі сфери, як міжнародний бізнес, де синхронний

переклад дає змогу проводити перемовини без залучення перекладачів. В освіті такі технології відкривають доступ до іноземних лекцій та онлайн-курсів, роблячи їх зрозумілими для студентів з різних країн. У сфері туризму синхронний переклад дозволяє мандрівникам легко орієнтуватися в чужій країні, а в службах підтримки підприємства можуть обслуговувати клієнтів різними мовами без додаткових витрат. У медицині системи голосового перекладу допомагають лікарям і пацієнтам долати мовний бар'єр у критичних ситуаціях, де швидкість розуміння є вирішальною [27].

Разом із тим існують важливі технічні виклики, які необхідно враховувати під час розробки таких систем. Одним з них є екосистема мобільних пристроїв, яка має обмежені обчислювальні ресурси. Нейронні мережі потребують значних потужностей, тому важливо правильно організувати роботу з Whisper: або частково запускати її локально, або використовувати хмарні сервери, оптимізуючи обсяг даних, що передаються мережею. Додатковим викликом є робота з нестабільним інтернет-з'єднанням, оскільки навіть невелика затримка може впливати на комфорт користувача. Тому необхідно застосовувати асинхронну обробку, кешування та паралельні потоки даних. Ще однією проблемою залишається природність синтезованого голосу. Сучасні TTS-системи значно покращилися, але й досі існують труднощі зі створенням голосів, які б мали природну інтонацію та відповідали контексту. Під час розробки важливо обирати такі моделі й налаштування, які забезпечують оптимальний баланс між якістю та швидкістю синтезу [21].

Об'єктом дослідження у цій роботі є сам процес автоматичного перекладу голосу у режимі реального часу, включаючи всі етапи обробки аудіо.

Предметом є методи та програмні засоби, які дозволяють реалізувати цей процес у вигляді мобільного застосунку. Наукова новизна роботи полягає у поєднанні кількох передових технологій: моделі Whisper для розпізнавання мовлення, Google Translate API для перекладу та хмарної інфраструктури

Firestore для збереження історії. Це поєднання дозволяє створити гнучку систему з низькою затримкою та високою точністю, яка адаптується до різних сценаріїв використання.

Практичне значення одержаних результатів створеного прототипу полягає в тому, що він може бути використаний у реальних умовах, служити основою для подальших досліджень або стати фундаментом для комерційного продукту. Така система може легко інтегруватися в онлайн-платформи, чат-боти, системи підтримки користувачів, туристичні сервіси, навчальні системи тощо. Моберекладач у реальному часі стає не просто програмою, а інструментом, який робить спілкування між людьми різних культур відкритішим і доступнішим.

Теоретичне значення одержаних результатів. Дослідження полягає в поглибленні наукових підходів до створення інтелектуальних систем синхронного голосового перекладу в реальному часі. Отримані результати розширюють сучасні уявлення про застосування нейронних мереж у задачах обробки природної мови, зокрема в контексті розпізнавання мовлення, машинного перекладу та синтезу голосу. У роботі узагальнено та систематизовано підходи до реалізації потокової обробки мовлення, що є важливим для подальшого розвитку систем автоматичної комунікації.

Апробація результатів кваліфікаційної роботи. Представлений програмний розроблений прототип застосунку, може використовуватись в синхронному голосовому перекладі у реальному часі. Отримані дані демонструють практичну придатність запропонованих методів та архітектурних рішень. Розроблена система може використовуватися, як основа для подальших досліджень у галузі автоматичного перекладу, а також для створення комерційних продуктів.

Структура роботи. Робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Загальний обсяг роботи становить 86 сторінок, включно з графічними матеріалами та програмним кодом.

