

AI Asystent do nauki języków

1. Wprowadzenie

Niniejszy dokument stanowi specyfikację wymagań (Software Requirements Specification – SRS) dla systemu Inteligentnego Asystenta Konwersacji Językowej. System ten ma na celu doskonalenie umiejętności językowych użytkowników w języku angielskim poprzez interaktywne konwersacje tekstowe i głosowe ze wspieranym przez AI asystentem.

2. Cele Projektu

Głównym celem projektu jest stworzenie samodzielnej aplikacji (mobilnej, webowej lub desktopowej), która:

1. Umożliwia naturalne konwersacje z Asystentem AI w języku angielskim.
2. Automatycznie koryguje błędy językowe użytkownika.
3. Obsługuje multimodalne wprowadzanie danych: tekst, głos (Speech-to-Text) i obrazy.
4. Zapewnia Asystentowi możliwość odpowiedzi w formie tekstowej i syntetyzowanej głosowo (Text-to-Speech).

3. Zespół Projektowy i Kompetencje

Członek Zespołu	Next.js, TS, React (Frontend)	UI/UX	Python (Backend)	AI	UML	Testy
Miłosz Ludwinek	Posiada	Posiada	Podstawy	Nie posiada	Podstawy	Nie posiada
Maciej Rybak	Podstawy	Nie posiada	Posiada	Nie posiada	Podstawy	Nie posiada
Filip Kucharski	Nie posiada	Podstawy	Posiada	Podstawy	Posiada	Podstawa
Michał Szczepański	Nie posiada	Nie posiada	Posiada	Nie posiada	Posiada	Podstawy

4. Struktura Danych (Model Komunikacji)

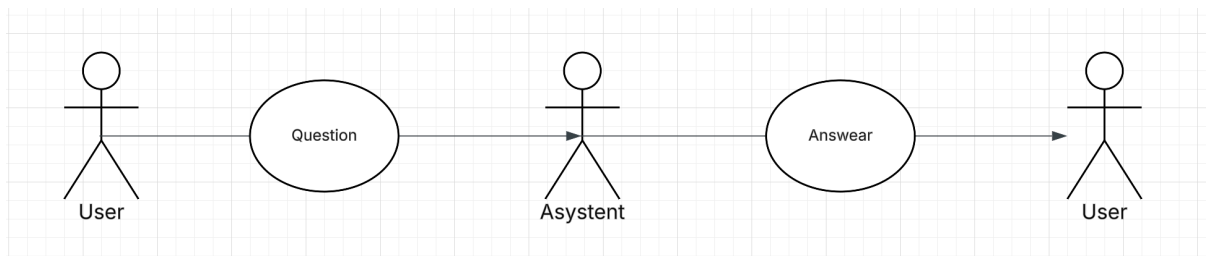
Ustalony Format Danych dla Komunikacji:

Obiekt	Pole	Typ Danych	Opis
Wiadomość Tekstowa/Głosowa	id	int (unique)	Unikalny identyfikator wiadomości.
	treść	string	Treść wiadomości tekstowej (po transkrypcji STT) lub odpowiedź Asystenta.
Wiadomość Obraz	id	int (unique)	Unikalny identyfikator obrazu w konwersacji.
	content	BLOB/base64	Zakodowana treść binarna pliku obrazu.

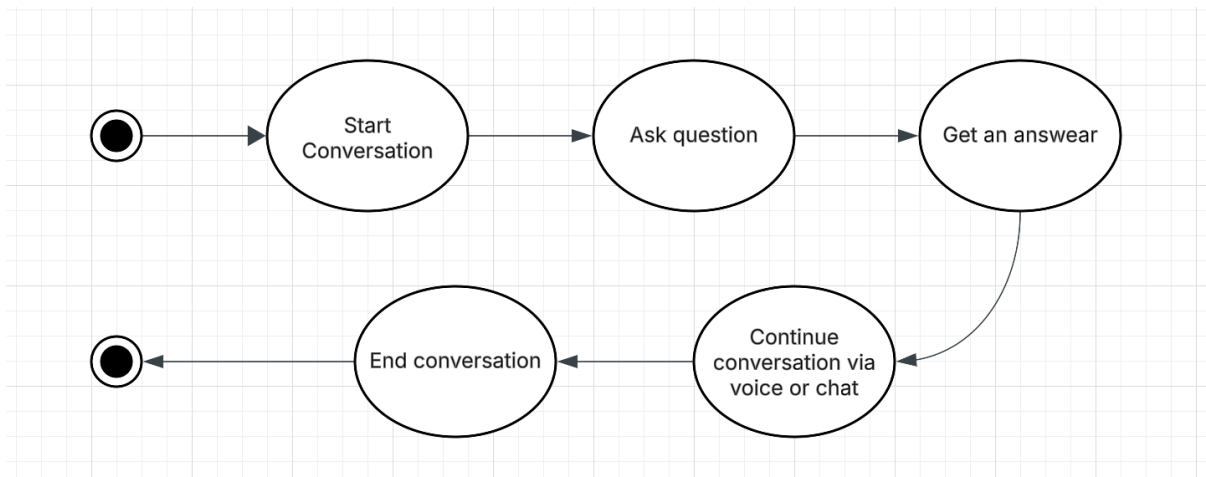
5. Modelowanie Systemu UML

Aktorzy	Użytkownik, Asystent
Opis	Użytkownik ma możliwość prowadzenia rozmów w celu doskonalenia umiejętności w języku angielskim. Użytkownik będzie miał możliwość wprowadzania wiadomości tekstowych, obrazów oraz wiadomości głosowych. Asystent będzie odpowiadał w języku angielskim i poprawiał błędne wiadomości użytkownika.
Dane	wiadomość/pytanie tekstowe/głosowe, obraz
Wyzwalacz	Po wprowadzeniu wiadomości przez użytkownika, asystent przetwarza i analizuje wprowadzoną treść
Odpowiedź	Asystent odpowiada na wprowadzoną odpowiedź w sposób tekstowy i głosowy
Uwagi	Odpowiedź głosowa jeśli użytkownik sobie tego zażyczy

5.1. Diagram Przypadków Użycia (Use Case Diagram)

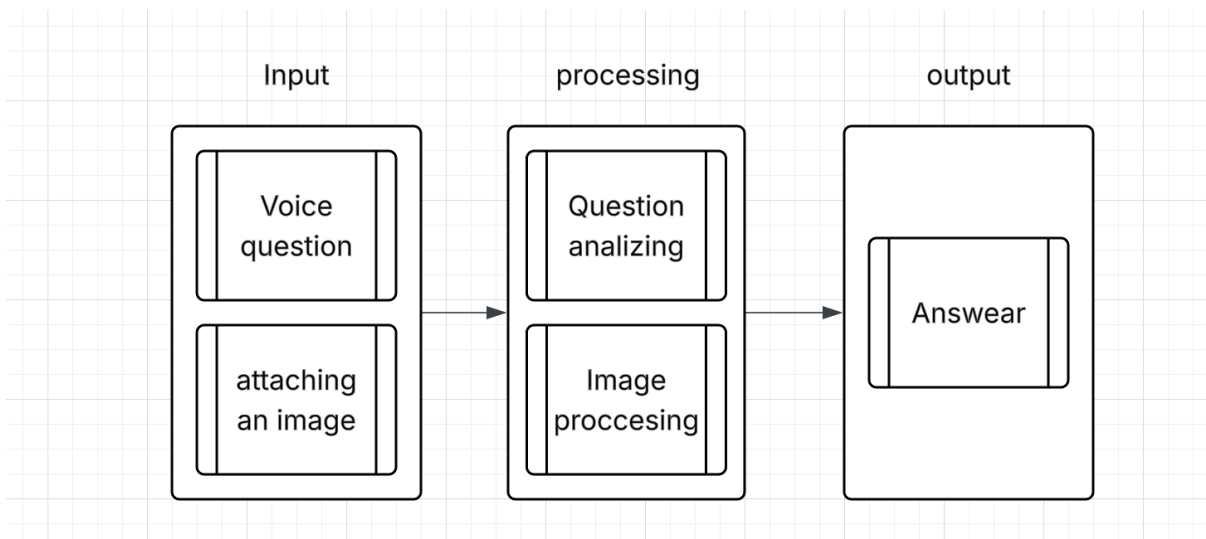


Aktorzy: Użytkownik, Asystent (System AI)



8. Proponowana Architektura Systemu

Proponuje się architekturę mikroservisową lub trójwarstwową w celu oddzielenia interfejsu (FE) od logiki biznesowej (BE) i usług AI, co zapewnia skalowalność i elastyczność w doborze technologii.



8.1. Warstwa Prezentacji (Frontend)

- **Technologia:** Next.js / React / TypeScript.

- **Odpowiedzialność:** Interfejs Użytkownika (UI), obsługa multimediów (kamera/mikrofon/upload), opcjonalne przełączanie interfejsu (PL/ENG), wyświetlanie konwersacji i poprawek.
- **Wdrożenie:** Aplikacja webowa (z możliwością konwersji na mobile/desktop dzięki React Native/Electron, jeśli jest to cel projektowy).

8.2. Warstwa Logiki Aplikacji (Backend - API Gateway / Serwer)

- **Technologia:** Python (np. z frameworkiem **FastAPI** lub **Django**).
- **Odpowiedzialność:**
 1. Autentykacja i zarządzanie sesjami (utrzymanie kontekstu 10 ostatnich wiadomości).
 2. Routing zapytań do odpowiednich modułów (NLP/Vision/TTS/STT).
 3. Interakcja z Bazą Danych (zapis/odczyt historii).
 4. Walidacja danych.

8.3. Warstwa Usług AI i Danych (AI Services & Persistence)

- **Usługi AI (Python/Microservices):**
 - **NLP Core:** Przetwarzanie wiadomości tekstowych, generowanie odpowiedzi (za pomocą dużego modelu językowego - LLM), korekta błędów gramatycznych i stylistycznych.
 - **Speech-to-Text (STT):** Konwersja wejścia głosowego na tekst (np. API od Google/AWS/OpenAI).
 - **Text-to-Speech (TTS):** Generowanie mowy z tekstu odpowiedzi Asystenta.
 - **Vision Core:** Analiza obrazów przesłanych przez użytkownika (np. LLM z multimodalnymi możliwościami).
- **Baza Danych:** Relacyjna (np. PostgreSQL) lub NoSQL (np. MongoDB) do przechowywania wiadomości (**id**, **treść**, **content**, **timestamp**, **user_id**, **session_id**).

Aplikacja do szlifowania umiejętności w języka angielskiego poprzez konwersacje z asystentem AI umożliwi użytkownikom konwersacje tekstowe i głosowe, a także wysyłanie zdjęć. Asystent, mający za zadanie poprawiać błędy, będzie odpowiadał na wiadomości użytkownika w języku angielskim. Rozmowy będą zapisywane, a projekt jest tworzony przez zespół z kompetencjami w Next.js, React, TypeScript i Pythonie.