

Relatório Técnico: Speech To Learn

Foco: Acessibilidade, Transcrição Automática e Avaliação do VLibras

1. Introdução

O projeto Speech To Learn é uma plataforma voltada para a educação inclusiva que integra o processamento de linguagem natural à interface de usuário. O objetivo central desta implementação é fornecer uma base sólida para a utilização do VLibras, permitindo que a transcrição gerada automaticamente sirva como insumo para a tradução em Libras, possibilitando futuras avaliações quantitativas sobre a eficiência da ferramenta no contexto de aprendizagem.

2. Arquitetura do Sistema

2.1 Backend (API de Transcrição)

O servidor atua como o motor de processamento, convertendo dados audiovisuais em dados textuais estruturados:

- **Tecnologia:** Construído em Node.js com Express e integrado ao modelo **Whisper da OpenAI**.
- **Fluxo de Dados:** O arquivo de vídeo é recebido via Multer, processado para extração de áudio, transcrita e os arquivos temporários são removidos para garantir a segurança e economia de armazenamento.
- **Saída Estruturada:** A API retorna um objeto JSON contendo o texto completo e um array de **segments** (contendo **start**, **end** e **text**), essenciais para a sincronização temporal no frontend.

2.2 Frontend (Client Side)

A interface foi projetada para ser um ambiente de testes e visualização:

- **Gerenciamento de Estado:** O componente **VideoPage** atua como orquestrador, sincronizando o tempo do vídeo (**currentTime**) com o texto exibido.
- **Interatividade:** Através do **Transcription.jsx**, o usuário pode navegar pelo vídeo clicando nos segmentos de texto, o que facilita a revisão de trechos específicos para avaliação.
- **Performance Local:** O uso de **URL.createObjectURL** permite que o vídeo seja reproduzido instantaneamente no navegador sem a latência de um upload completo para um servidor de streaming.

3. Acessibilidade e VLibras

O diferencial deste projeto é a camada de acessibilidade desenhada para o usuário surdo ou com baixa visão:

- **VLibras:** O projeto implementa o widget do VLibras para tradução automática da transcrição, permitindo que pesquisadores e usuários avaliem quantitativamente a precisão da tradução sinalizada em relação ao texto original gerado pelo Whisper.

- **Alto Contraste:** Sistema de cores dinâmico (Fundo preto / Texto amarelo) otimizado para usuários com sensibilidade visual ou baixa visão.
- **Feedback Visual e Navegação:** Botões e estados de foco são reforçados para garantir naveabilidade por teclado e leitura clara de tela.

4. Guia de Operação (Docker)

Para garantir a reproduzibilidade dos testes de eficiência, o ambiente é containerizado:

- **Subir Ambiente:** make up.
- **Encerrar Ambiente:** make down.