Relatório Final: API de Processamento de Áudio

Danilo Chagas Clemente, Lara Ramos Linhares, Lucas Alvarenga Lopes, Marco Túlio Amaral

February 10, 2025

1 Introdução

O processamento de áudio baseado em inteligência artificial tem se tornado uma tecnologia essencial para diversas aplicações, como acessibilidade, atendimento automatizado e análise de discursos. No entanto, esse tipo de sistema envolve riscos associados à proteção de dados pessoais, especialmente quando lida com áudios sensíveis. Este projeto propõe o desenvolvimento de uma API para transcrição, processamento de texto e conversão para áudio, garantindo a segurança e a conformidade com a Lei Geral de Proteção de Dados (LGPD).

A escolha desse problema se justifica pela necessidade da *Hongik Uniersity* de um sistema de processamento de áudio que funcionasse para responder as dúvidas dos visitantes do Museu Nacional Da Coréia. Nesse caso, os visitantes fariam perguntas com a própria voz, e teriam suas perguntas respondidas em forma de áudio.

2 Desenvolvimento

A API desenvolvida realiza três principais funções:

- Transcrição de áudio usando o *OpenAI Whisper*.
- Processamento do texto transcrito com o modelo *Ollama*.
- Conversão do texto gerado em áudio utilizando a *API* do *ElevenLabs*.

A implementação e os códigos deste projeto estão disponíveis no seguinte repositório do GitHub:

`https://github.com/Am4ral/api-sd`

3 Relatório de Impacto à Proteção de Dados Pessoais (RIPD)

3.1 Metodologia de Análise de Riscos

A análise de riscos desta API foi realizada utilizando a modelagem de ameaças STRIDE e a abordagem de Torr (2005). O STRIDE categoriza ameaças em seis grupos principais:

- **Spoofing (Falsificação de identidade)**
- **Tampering (Alteração de dados)**
- **Repudiation (Repúdio de ações)**
- **Information Disclosure (Divulgação não autorizada de informações)**
- **Denial of Service - DoS (Negação de serviço)**
- **Elevation of Privilege (Elevação de privilégios)**

3.2 Identificação e Avaliação dos Riscos

Com base no modelo STRIDE, foram identificados os seguintes riscos para cada serviço da API:

3.2.1 WhisperAI - Transcrição de Áudio

- **Risco 1 - Divulgação de Informações (Information Disclosure)**: O áudio pode conter dados pessoais sensíveis, e sua transcrição pode ser armazenada sem criptografia.
- **Risco 2 - Negação de Serviço (Denial of Service)**: Um alto volume de requisições pode sobrecarregar o serviço de transcrição.

3.2.2 Ollama - Processamento de Texto

- **Risco 3 - Alteração de Dados (Tampering)**: Um atacante pode modificar a resposta gerada pelo modelo antes de ser enviada ao usuário.
- **Risco 4 - Repúdio de Ações (Repudiation)**: A API não registra logs de quem enviou solicitações, impossibilitando auditoria adequada.

3.2.3 ElevenLabs - Conversão de Texto em Áudio

- **Risco 5 - Falsificação de Identidade (Spoofing)**: Um atacante pode tentar gerar áudios falsos se o sistema não tiver restrição de acesso.

3.3 Medidas para Mitigar os Riscos

Para cada risco identificado, foram propostas as seguintes medidas:

3.3.1 Medidas para WhisperAI

- **Criptografia** dos arquivos de áudio e transcrições antes do armazenamento.
- **Rate limiting** para impedir ataques de negação de serviço.

3.3.2 Medidas para Ollama

- **Assinatura digital** das respostas geradas pelo modelo para garantir sua integridade.
- **Implementação de logs e auditoria** para rastreamento de acessos.

3.3.3 Medidas para ElevenLabs

- **Autenticação e Autorização**: Controle de acesso baseado em autenticação para garantir que apenas usuários autorizados possam fazer requisições ao serviço.
- **Monitoramento de Requisições**: Implementação de logs e ferramentas de monitoramento para detectar e responder a padrões de uso suspeitos.

4 Considerações Finais

A API desenvolvida mostrou-se eficiente na execução das três etapas principais: transcrição, processamento e conversão de áudio. Entretanto, houveram problemas com o tratamento das respostas do modelo *Ollama*, o que prejudicou o funcionamento da API. A adoção de medidas de segurança e privacidade foi fundamental para garantir a conformidade com a LGPD e mitigar riscos associados ao uso de dados pessoais. A modelagem de ameaças utilizando STRIDE permitiu a identificação de vulnerabilidades críticas e a aplicação de soluções adequadas.