



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

Arthur de Amorim Macena

ANNOTATION TOOL: ANOTAÇÕES DE DADOS POR MEIO DE ÁUDIOS

CAMPINA GRANDE - PB

2023

Arthur de Amorim Macena

ANNOTATION TOOL: ANOTAÇÕES DE DADOS POR MEIO DE ÁUDIOS

Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

Orientador : Cláudio Elízio Calazans Campelo

CAMPINA GRANDE - PB

2023

Arthur de Amorim Macena

ANNOTATION TOOL: ANOTAÇÕES DE DADOS POR MEIO DE ÁUDIOS

Trabalho de Conclusão Curso apresentado ao Curso Bacharelado em Ciência da Computação do Centro de Engenharia Elétrica e Informática da Universidade Federal de Campina Grande, como requisito parcial para obtenção do título de Bacharel em Ciência da Computação.

BANCA EXAMINADORA:

**Cláudio Elízio Calazans Campelo
Orientador – UASC/CEEI/UFCG**

**Carlos Eduardo Santos Pires
Examinador – UASC/CEEI/UFCG**

**Francisco Vilar Brasileiro
Professor da Disciplina TCC – UASC/CEEI/UFCG**

Trabalho aprovado em: 28 de Junho de 2023.

CAMPINA GRANDE - PB

RESUMO

Este trabalho de conclusão de curso aborda o desenvolvimento de uma ferramenta *web* para o projeto de anotação de dados do Laboratório de Computação Inteligente Aplicada (LACINA). Essa ferramenta tem como foco anotações de áudio baseadas em diálogos, com o objetivo de facilitar a criação de conjuntos de dados obtidos para o treinamento de sistemas de diálogo baseados em *machine learning*. A ferramenta foi projetada para fornecer uma plataforma eficiente e intuitiva, permitindo a identificação de falantes e a marcação de intenções e entidades nos diálogos em áudio. O projeto destaca a importância das anotações de áudio baseadas em diálogos para o avanço dos sistemas de diálogo, e discute as vantagens de se ter conjuntos de dados anotados, bem como as considerações éticas e de privacidade associadas a esse processo. A sessão de desenvolvimento abordou os aspectos técnicos da ferramenta, incluindo a escolha das tecnologias e *frameworks* utilizados, os desafios enfrentados durante o andamento e as soluções adotadas para superá-los. Foram destacadas as funcionalidades desejáveis, como uma interface de usuário amigável, recursos avançados de reprodução e visualização sincronizada de áudio e texto anotados. Em seguida, foram discutidas as aplicações práticas da ferramenta no contexto de projetos de *machine learning* e processamento de linguagem natural. Os conjuntos de dados anotados criados com a ferramenta podem ser usados para treinar modelos de reconhecimento de fala e compreensão de linguagem, contribuindo para o desenvolvimento de sistemas de diálogo conversacional. Este trabalho visa fornecer uma visão abrangente desta área em constante evolução, destacando seu impacto e seu potencial para impulsionar a pesquisa e o desenvolvimento de aplicações de *machine learning* baseadas em áudio. A ferramenta *web* de anotações de áudio baseadas em diálogos representa uma contribuição significativa para o campo do processamento de linguagem natural e sistemas de diálogo, facilitando a criação de conjuntos de dados anotados de alta qualidade e impulsionando o avanço e o desenvolvimento de sistemas de diálogo mais eficientes, naturais e precisos.

ANNOTATION TOOL: ANOTAÇÕES DE DADOS POR MEIO DE ÁUDIOS

ABSTRACT

This course completion work addresses the development of a web tool for the data annotation project at the Computer Laboratory Intelligent Applied (LACINA). This tool focuses on dialogue-based audio annotations, with the aim of facilitating the creation of datasets obtained for training dialogue systems based on machine learning. The tool was designed to provide an efficient and intuitive platform, allowing the identification of speakers and the marking of intentions and entities in audio dialogues. The project highlights the importance of dialog-based audio annotations to the advancement of dialog systems, and discusses the advantages of having annotated datasets, as well as the ethical and privacy considerations associated with this process. The development session addresses the technical aspects of the tool, including the choice of technologies and frameworks used, the challenges faced during the course and the solutions adopted to overcome them. Desirable functionalities were highlighted, such as a friendly user interface, advanced playback features and synchronized visualization of audio and annotated text. Then, practical applications of the tool in the context of machine learning and natural language processing projects were discussed. The annotated datasets created with the tool can be used to train speech recognition and language comprehension models, contributing to the development of conversational dialogue systems. This work aims to provide a comprehensive overview of this ever-evolving area, highlighting its impact and its potential to drive research and development of audio-based machine learning applications. The dialog-based audio annotations web tool represents a significant contribution to the field of natural language processing and dialog systems, facilitating the creation of high quality annotated datasets and driving the advancement and development of more advanced dialog systems. efficient, natural and accurate.

Annotation Tool: Anotações de dados por meio de áudios

Trabalho de Conclusão de Curso

Arthur de Amorim Macena (Aluno), Cláudio Campelo (Orientador)

Departamento de Sistemas e Computação
Universidade Federal de Campina Grande
Campina Grande, Paraíba - Brasil

RESUMO

Este trabalho de conclusão de curso aborda o desenvolvimento de uma ferramenta *web* para o projeto de anotação de dados do Laboratório de Computação Inteligente Aplicada (LACINA). Essa ferramenta tem como foco anotações de áudio baseadas em diálogos, com o objetivo de facilitar a criação de conjuntos de dados obtidos para o treinamento de sistemas de diálogo baseados em *machine learning*. A ferramenta foi projetada para fornecer uma plataforma eficiente e intuitiva, permitindo a identificação de falantes e a marcação de intenções e entidades nos diálogos em áudio. O projeto destaca a importância das anotações de áudio baseadas em diálogos para o avanço dos sistemas de diálogo, e discute as vantagens de se ter conjuntos de dados anotados, bem como as considerações éticas e de privacidade associadas a esse processo. A sessão de desenvolvimento aborda os aspectos técnicos da ferramenta, incluindo a escolha das tecnologias e *frameworks* utilizados, os desafios enfrentados durante o andamento e as soluções adotadas para superá-los. Foram destacadas as funcionalidades desejáveis, como uma interface de usuário amigável, recursos avançados de reprodução e visualização sincronizada de áudio e texto anotado. Em seguida, foram discutidas as aplicações práticas da ferramenta no contexto de projetos de *machine learning* e processamento de linguagem natural. Os conjuntos de dados anotados criados com a ferramenta podem ser usados para treinar modelos de reconhecimento de fala e compreensão de linguagem, contribuindo para o desenvolvimento de sistemas de diálogo conversacional. Este trabalho visa fornecer uma visão abrangente dessa área em constante evolução, destacando seu impacto e seu potencial para impulsionar a pesquisa e o desenvolvimento de aplicações de *machine learning* baseadas em áudio. A ferramenta *web* de anotações de áudio baseadas em diálogos representa uma contribuição significativa para o campo do processamento de linguagem natural e sistemas de diálogo, facilitando a criação de conjuntos de dados anotados de alta qualidade e impulsionando o avanço e o desenvolvimento de sistemas de diálogo mais eficientes, naturais e precisos.

PALAVRAS-CHAVE

Áudio, Anotação, Desenvolvimento *Web*, *Machine Learning*

Os autores retêm os direitos, ao abrigo de uma licença Creative Commons Atribuição CC BY, sobre todo o conteúdo deste artigo (incluindo todos os elementos que possam conter, tais como figuras, desenhos, tabelas), bem como sobre todos os materiais produzidos pelos autores que estejam relacionados ao trabalho relatado e que estejam referenciados no artigo (tais como códigos fonte e bases de dados). Essa licença permite que outros distribuam, adaptem e evoluam seu trabalho, mesmo comercialmente, desde que os autores sejam creditados pela criação original.

1 INTRODUÇÃO

O desenvolvimento de ferramentas *web* para anotações de áudio baseadas em diálogos tem se tornado cada vez mais relevante no contexto de projetos de *machine learning* e processamento de linguagem natural. Essas ferramentas desempenham um papel fundamental na coleta e na organização de dados de diálogos em áudio, permitindo a criação de conjuntos de dados anotados essenciais para treinar e aprimorar modelos de sistemas de diálogo. Tais *frameworks* envolvem a atribuição de rótulos, transcrições ou marcações temporais a segmentos de áudio específicos, permitindo que os algoritmos de aprendizado de máquina compreendam e extraiam informações relevantes dos diálogos.

Este trabalho se concentra no desenvolvimento de uma ferramenta para um projeto específico do Laboratório de Computação Inteligente Aplicada (LACINA) de anotação de dados, explorando as possibilidades e os desafios inerentes a essa abordagem. O objetivo principal é fornecer uma plataforma eficiente e intuitiva que facilite o processo de anotação de diálogos, a transcrição de áudio, a identificação de falantes e a marcação de intenções e entidades.

No presente trabalho, serão apresentados os fundamentos e a importância das anotações de áudio baseadas em diálogos. Serão discutidas as vantagens de se ter um conjunto de dados anotado para treinar modelos de sistemas de diálogo, bem como as limitações e os desafios enfrentados nesse processo.

Serão exploradas as principais funcionalidades e características desejáveis em uma ferramenta *web* de anotações de áudio [4]. Isso inclui a interface de usuário amigável, que facilite a interação e a navegação pelos áudios e pelas anotações, além de recursos avançados, como a capacidade de adicionar anotações temporais precisas.

O objetivo do presente trabalho será discutir o processo de desenvolvimento da ferramenta, desde a escolha das tecnologias e *frameworks* adequados até a implementação das funcionalidades específicas, como a disposição da ferramenta para diálogos. Serão apresentados os desafios encontrados durante o desenvolvimento e as soluções adotadas para superá-los, levando em consideração aspectos como segurança, escalabilidade e desempenho. Além disso, serão discutidas as aplicações práticas da ferramenta de anotações de áudio baseadas em diálogos, explorando o seu potencial em projetos de pesquisa e desenvolvimento de sistemas de diálogo e possibilitando a criação de conjuntos de dados anotados para treinamento de modelos.

Por fim, serão abordados os desafios e as perspectivas futuras relacionadas à ferramenta de anotações de áudio. Serão discutidas possíveis melhorias e expansões, como a integração com outras ferramentas de processamento de linguagem natural, a incorporação de técnicas de aprendizado ativo para otimizar o processo de

anotação e a adaptação da ferramenta para diferentes idiomas e contextos específicos.

Em resumo, este trabalho tem como objetivo principal o desenvolvimento de uma ferramenta *web* de anotações de áudio baseadas em diálogos, visando facilitar e agilizar o processo de criação de conjuntos de dados anotados para treinamento de modelos de sistemas de diálogo. Ao explorar as funcionalidades da ferramenta, bem como suas vantagens e aplicações, espera-se contribuir para o avanço e aprimoramento de sistemas de diálogo baseados em *machine learning*.

2 PROBLEMA E SOLUÇÃO

Nesta seção, expõe-se a motivação da anotação de áudio no projeto, a estrutura da ferramenta para atingir o proposto e os detalhes técnicos envolvidos no trabalho.

2.1 Anotação de Áudio vs Anotação de Texto

As anotações de dados podem ser feitas de diferentes maneiras. Uma das problemáticas abordadas é a discussão acerca das anotações de áudio se fazerem necessárias e o questionamento sobre as vantagens trazidas por elas, visto que as anotações de texto são de mais fácil armazenamento e de mais fácil implementação (são comumente utilizadas em aplicações desse contexto). Respondendo a problemática, as principais motivações são: as pessoas são contratadas para uma demanda de texto, mas podem nunca ter trabalhado com essa estrutura e ficam passíveis de erros, além do fato das anotações de áudio terem vantagens como facilitação das contextualizações, aplicação com base em diferentes domínios e tem como base a captura de nuance, uma vez que não é possível obtê-las através do texto. Dessa forma, ela se faz bastante útil, ampliando os potenciais do *software* e o tornando mais robusto.

2.2 Interface da Ferramenta

Uma interface *web* para anotações de áudio pode enfrentar uma série de problemas que podem afetar a eficiência e a experiência do usuário. Alguns dos problemas comuns incluem a Usabilidade e Interface de Usuário (UI). Nessa linha de pensamento, é apresentada uma complexidade excessiva, que dificulta a utilização da ferramenta, como a falta de controle para o mecanismo de áudios. Outra dificuldade comum nesse âmbito é a má disposição de elementos, tornando o *software* não intuitivo, podendo confundir a forma de utilizá-lo. Por último, a colaboração é um problema bastante grave quando se trata da estrutura da ferramenta, pois em cenários colaborativos (quando várias pessoas estão anotando o mesmo áudio) pode haver conflitos (desacordo) nas anotações. Portanto, é necessário um mecanismo, como processamento de dados progressivos, para resolver esses conflitos e manter a consistência nos dados anotados. Diante disso, o foco definido foi desenvolver uma interface de usuário intuitiva e amigável, com uma disposição clara dos elementos e *feedback* adequado para orientar os usuários e utilizar algoritmos de sincronização para garantir a precisão na correspondência entre o áudio e o texto anotado.

2.3 Aspectos Operacionais

Durante o processo de busca técnica para compor a ferramenta, foram identificados problemas relacionados à manipulação de áudio,

gerenciamento de grandes arquivos e armazenamento de anotações. Para resolver o problema de manipulação de áudio, foi decidido utilizar uma biblioteca de áudio para aproveitar melhor a oferta de recursos avançados para reprodução e controle do áudio. Isso permitiu uma reprodução suave e precisa do áudio, sincronizada com as anotações. No que diz respeito aos grandes arquivos de áudio, adotou-se técnicas de otimização, como o carregamento progressivo ou a segmentação do áudio, a fim de melhorar o desempenho e a eficiência do processo de reprodução. Para o armazenamento e gerenciamento de anotações, utilizou-se um banco de dados para manter as anotações dos usuários e foi implementado um sistema de autenticação e autorização para controle de acesso. Essas soluções contribuíram para o desenvolvimento de uma interface *web* eficiente e funcional para anotação de áudio em diálogos, tornando possível o armazenamento de anotações.

3 ARQUITETURA DA SOLUÇÃO

Esta seção fornece detalhes sobre a arquitetura e os fundamentos teóricos utilizados para a construção da ferramenta. Ela é dividida em: Componentes do sistema, Primeira etapa da configuração, Componentes da ferramenta, Gerenciamento de dados e Configuração Final da Ferramenta. Inicialmente, são fornecidos detalhes sobre as tecnologias utilizadas no projeto e, em seguida, é descrito como funciona o fluxo da informação da aplicação. Após isto, é descrito como pode-se realizar a configuração da ferramenta, a fim de que se torne útil para o tema de anotação de dados em áudio, e como são feitas as marcações temporais e de texto. Depois, é descrito como funcionam os componentes utilizados para melhorar a estilização da ferramenta, com o objetivo de se tornar uma interface de mais fácil uso para o usuário. Além disso, é descrito como foram feitos os ajustes de dados para que possam ser armazenados e devidamente gerenciados para qualquer fim que as informações de anotação se façam necessárias. E, por fim, tem-se uma seção dedicada a explicar como é feito para tornar a ferramenta útil não só para anotação de áudios, mas também para o outro foco do projeto que são os diálogos.

3.1 Componentes do Sistema

O sistema utilizado para desenvolvimento da ferramenta é o *software Annotation Tool*. Ele é uma aplicação *web* implementada com tecnologias como *JavaScript*¹, no *backend* e no *frontend*, mas utilizando no *frontend* o *framework React*², no *backend* o *framework* utilizado é o *NodeJS*³, e são modularizados em contêineres com o *Docker*⁴, para levantar a aplicação, além de contêineres para aplicação, tem-se para o banco de dados também que é o *MongoDB*⁵.

A lógica de negócio do sistema foi construída inteiramente em *JavaScript*, pela sua popularidade, e as escolhas mais populares para *frontend* e *backend* utilizando essa linguagem são respectivamente *React* e *NodeJS*, pela sua diversidade de bibliotecas e pela curva de aprendizagem alta para facilitar novos desenvolvedores a entenderem o sistema. A modularização com o *Docker* se dá para melhor

¹<https://www.javascript.com>

²<https://react.dev>

³<https://nodejs.org/en>

⁴<https://www.docker.com/>

⁵<https://www.mongodb.com>

utilização do sistema, visto que é feito dependendo de uma configuração única dos dois casos e podem funcionar simultaneamente, mesmo que de forma local. O MongoDB é um banco não relacional utilizado no projeto que possui uma ótima integração com o *framework NodeJS*, pois eles tem uma sintaxe muito parecidas e tornam a programação mais coesa. Os bancos não relacionais, de acordo com Chodorow[1], são bancos que tem facilidade de uso e escalabilidade horizontal. Diante das tecnologias, o fluxo da aplicação consiste em administradores da aplicação que cadastram usuários e esses são responsáveis por, dentre várias coisas, cadastrarem novos diálogos para outros usuários poderem fazer as anotações deles, sendo todos os dados guardados no banco em tempo de execução do sistema. Para o presente trabalho é proposto que os diálogos possam ser salvos também em formatos de áudio e as anotações aconteçam também em arquivos desse formato, não apenas em texto.

Com esse objetivo, foi utilizado a biblioteca do *React* para manipulação de áudio *wavesurfer.js*[2], pois é uma biblioteca¹ que oferece uma boa integração com o *frontend* e fornece também uma boa gama de recursos que podem ser usados, tais como: registro de tempos, recursos de play, pause e zoom para maior controle das funções necessárias para marcação de dados, além de conseguir carregar recursos de áudio e ter uma interface agradável para o uso. O fluxo para essa nova ferramenta será o mesmo do anterior, porém o usuário agora além de poder de criar diálogos de texto, poderá criar diálogos por áudio e assim outro usuário cadastrado poderá fazer as anotações pelos arquivos de áudio.

3.2 Primeira Etapa de Configuração

Essa seção conta como foi feito o passo a passo para a configuração da ferramenta de anotação de áudio. A principal parte foi adicionar a biblioteca *wavesurfer.js* no projeto, através do recurso do *yarn*², e criar um novo componente para a nova estrutura, mantendo os padrões de uso de código, facilitando para futuras manutenções, e mantendo a qualidade do código como mencionado por Martin[3]. Em seguida está um exemplo de criação do componente:

```
import WaveSurfer from 'wavesurfer.js'

const wavesurfer = WaveSurfer.create({
  container: document.body,
  waveColor: 'rgb(200, 0, 200)',
  progressColor: 'rgb(100, 0, 100)',
  url: '/examples/audio/audio.wav',
})

wavesurfer.once('interaction', () => {
  wavesurfer.play()
})
```

Acima, é possível ver um exemplo de criação de estrutura para o *wavesurfer*, no qual é definido o endereço do arquivo de áudio a ser utilizado para demonstração, o contêiner que será posto a faixa de áudio, a cor tanto da barra, quanto do progresso feito, além da estrutura de *play* para ativar o áudio assim que tiver alguma interação na ferramenta.

3.3 Componentes da Ferramenta

Para atender a todo o grau de utilização da aplicação, é necessário adicionar ferramentas para que o uso seja mais dinâmico e o usuário tenha uma melhor experiência. Para isso foi necessário adicionar botões de *play/pause* e botão de zoom, para facilitar a marcação de dados no tempo mais correto possível, e, além dos componentes mais importantes para o trabalho, o componente de disponibilização de tempo do áudio e o de marcação da faixa de áudio. Para os botões foi utilizada uma biblioteca de estilização do *React Antd*³, assim como para o botão de controle de zoom. Já para os de controle de tempo da faixa de áudio e o de marcação foi utilizado um processo diferente. Para essa tecnologia, foram instalados os plugins da biblioteca do *wavesurfer* e para o controle de tempo foi instalado o plugin de *TimeLinePlugin*⁴. Já para o de marcação de dados o *RegionsPlugin*⁵, abaixo uma imagem de exemplo para ilustração:

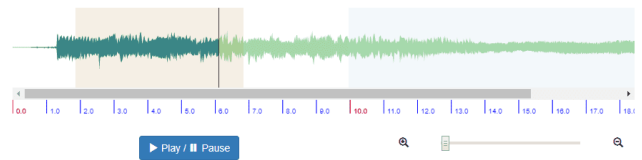


Figura 1: Ilustração da Ferramenta

3.4 Gerenciamento de dados

A etapa de armazenamento de dados de anotações de áudios em um banco de dados não relacional, como o *MongoDB*, é fundamental para garantir a eficiência e escalabilidade do sistema de anotações.

Nesse processo, é necessário estabelecer a conexão com o *MongoDB*, utilizando uma biblioteca adequada para a linguagem de programação escolhida, no caso do *node* foi utilizada a biblioteca *mongoose*. Essa conexão permitirá a interação com o banco de dados, viabilizando operações de armazenamento, recuperação e manipulação dos dados e a etapa de configuração é possível ser vista na figura abaixo:

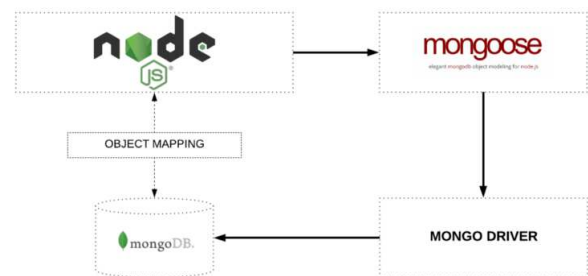


Figura 2: Comunicação do *backend* com o banco

¹<https://coodesh.com/blog/dicionario/o-que-e-biblioteca/>

²<https://yarnpkg.com/package/react>

³<https://ant.design>

⁴<https://wavesurfer-js.org/example/timeline/>

⁵<https://wavesurfer-js.org/example/regions/>

Para estruturar os dados das anotações, é preciso definir o esquema que especifica os campos e seus tipos. No caso das anotações de áudio, o esquema pode incluir informações como trecho de áudio, texto da anotação, usuário responsável e outras informações relevantes. O *MongoDB* oferece flexibilidade nesse aspecto, permitindo que diferentes documentos tenham campos distintos.

Após a definição do esquema, são criados modelos de dados que representam as anotações de áudio. Esses modelos facilitam as operações CRUD, fornecendo uma interface para interagir com o banco de dados de forma simplificada. Por meio desses modelos, é possível criar, ler, atualizar e excluir as anotações.

Durante o processo de anotação, os dados são coletados e formatados de acordo com o esquema definido. Por exemplo, quando um usuário adiciona uma nova anotação de áudio, os campos relevantes, como o trecho de áudio e o texto da anotação, são armazenados no documento correspondente. Em seguida, esse documento é salvo no *MongoDB* como uma nova entrada na coleção de anotações.

A recuperação dos dados também é uma etapa importante. É possível realizar consultas ao *MongoDB* para obter as anotações de áudio desejadas, utilizando critérios de busca como o ID do diálogo ou informações relacionadas. As consultas podem ser filtradas e ordenadas conforme necessário, permitindo uma recuperação eficiente dos dados.

Além disso, o *MongoDB* oferece recursos para atualização e exclusão dos documentos de anotação, possibilitando que os usuários modifiquem ou removam suas anotações quando necessário. Isso mantém o banco de dados atualizado com as alterações feitas pelos usuários.

Portanto, é importante destacar que o *MongoDB* também suporta o armazenamento de arquivos, como os próprios áudios dos diálogos. Isso permite que os diálogos sejam salvos como arquivos de áudio, complementando a capacidade de anotações em formato de texto. Dessa forma, a plataforma oferece suporte tanto para armazenar as anotações em texto quanto para associá-las aos arquivos de áudio correspondentes.

3.5 Configuração Final da Ferramenta

Para a etapa final de configuração da ferramenta, é necessário que os áudios estejam organizados em uma estrutura de dados para a disponibilização em forma de diálogo. Para isso, foi usado a criação de um componente que armazenasse os arquivos de áudio em forma de lista, tornando possível, através dos atributos de cada arquivo de áudio obtido do banco de dados e retornado do *backend*, o alinhamento e a ordenação, bem como trazer as informações já adicionadas.

4 ESTUDO DE CASO

Esta seção apresenta um estudo de caso, que tem como objetivo demonstrar as possibilidades de utilização da ferramenta *Annotation Tool*. Com o intuito de enaltecer o trabalho, será feita uma demonstração com mais detalhes do funcionamento da ferramenta, com uma visão mais aprofundada de seu uso e de como é disposta sua usabilidade. Após isso, serão apresentadas as considerações desse cenário.

4.1 Cenário do Estudo de Caso

Vamos imaginar o seguinte cenário: Um usuário quer criar um novo diálogo em formato de áudio, cujo o principal objetivo é que outro usuário pegue esse diálogo para fazer a anotação de dados, com isso, para atender aos requisitos funcionais da ferramenta o usuário vai para etapa de cadastro, no qual vê um form para identificação de campos como Cliente e Agente, um campo para selecionar a linguagem do áudio outro para upload do áudio e ele vai preenchendo os campos até que o diálogo esteja completo e ele tenha feito o upload de cada parte do diálogo entre o Cliente e o Agente, após essa perspectiva ele salva o diálogo. O diálogo recém criado vai para uma nova lista de diálogos de áudios, onde os usuários podem ver para checagem de dados, excluir, ou ir para ação de marcação de dados utilizando a *Anotação de Áudio*.

4.2 Utilizando a Anotação de Áudio

Para fazer a anotação de áudio, um usuário que não foi o criador do diálogo vai na listagem de anotações de áudio e aperta o botão de Anotação. Ao apertar o botão, o usuário é direcionado para outra página com os áudios inseridos e na mesma posição de inserção, e com os áudios amostra a esquerda com a demonstração dos áudios do Agente e a direita os áudios do Cliente, com identificadores de tempo, além de botões de *play/pause* e zoom para maior autonomia em cada parte do diálogo. Para realizar a anotação o usuário clica no *play* do áudio que ele deseja anotar que começa a tocar e a correr uma palheta de identificação do tempo em que o áudio se encontra, caso deseje fazer a anotação o usuário segura o botão esquerdo do mouse no tempo inicial que deseja fazer a marcação e arrastando até o tempo final que deseja, desde que esteja dentro dos limites do tempo do áudio, soltando o botão esquerdo, cria um campo de marcação onde o usuário poderá apertar com dois cliques e abrirá uma *pop up* de *labels* que estará todas as *labels* registradas no banco de dados do *backend*, escolhendo uma *label* para a marcação de dados, ela fica salvo abaixo do áudio essa seção com o nome da *label* e um id dela para aquele áudio e na marcação fica a cor que está salvo a *label* no banco de dados. Abaixo do áudio ficam as anotações feitas no áudio e clicando em qualquer uma anotação ele leva a palheta para o início dessa seção e dará o *play* nela.

4.3 Aspectos Favoráveis da Ferramenta

Diante dos testes realizados, podemos constatar que a ferramenta é bastante simples e intuitiva, seguindo os padrões do projeto. Além disso, ela se mostra extremamente eficiente no gerenciamento de volume de informações. Mesmo quando há uma grande carga de dados, a ferramenta não interferirá no diálogo apresentado, uma vez que as informações são contidas dentro do áudio e apresentadas apenas na área de visualização, não causando dificuldades na marcação de dados. Mesmo que os áudios sejam longos, há a opção de ajustar o tamanho necessário para tornar a marcação mais confortável. Além disso, foi observado que a ferramenta transmite uma maior quantidade de informações em comparação com a ferramenta de anotação de texto. Isso se deve ao fato de que, no áudio, é possível capturar o conteúdo do ambiente ao redor da conversa, o que amplia ainda mais a facilidade, compreensão e revisão do conteúdo. Por fim, a ferramenta oferece informações mais confiáveis, uma vez que elimina o risco de erros de digitação e permite

a preservação exata de elementos falados, como citações ou dados numéricos.

5 DISCUSSÃO E TRABALHOS FUTUROS

Esta seção discute as limitações da aplicação desenvolvida e aponta para trabalhos futuros.

Este trabalho apresenta uma ferramenta desenvolvida para o projeto de anotação de dados do Laboratório de Computação Inteligente Aplicada (LACINA), com o objetivo de torná-lo um projeto ainda mais robusto e com uma maior gama de funções. Com a anotação de áudio, ele se torna capaz de ser utilizado para análise de sentimentos com base em uma forma muito mais simples: a partir dos identificadores de fala que forem selecionadas. Dessa forma, a aplicação é uma boa fonte de dados para treinamento de Inteligências Artificiais (I.A.) e com mais facilidade para anotações, já que grandes diálogos textuais podem ser reduzidos a alguns áudios que tem uma anotação muito precisa com base no tempo. Para limitações da ferramenta, uma delas que vou abordar é a não existência de transcrição de dados, visto que, o armazenamento de áudios, podem se tornar bastante pesados ao longo do tempo, essa limitação peca um pouco no escalonamento do projeto. Além dessa, temos que os áudios não são analisados antes de serem processados, portanto, não temos garantia de áudios bem definidas. Além disso, a ferramenta tem dificuldades de trabalhar quando a versão do *nodeJS* se encontra desatualizada, tendo versões específicas para seu funcionamento em versões antigas. Para trabalhos futuros podem ser vistas formas de transcrição para os áudios, de forma que o armazenamento se torne menos custoso e uma forma mais fácil de visualização dos dados. Além disso, pode ser vista uma forma de garantir a qualidade dos áudios, para que seja mais fácil a anotação manual realizada, e evitar a variedade de qualidade na gravação, o que também auxiliaria na transcrição, seja ela de qualquer forma. Portanto, a ferramenta se torna bastante útil no contexto do projeto complementando ao que já se fazia existente e aumentando a amplitude do que se pode ser feito pelos usuários.

6 AGRADECIMENTOS

Gostaria de agradecer ao corpo docente do curso de Ciência da Computação da UFCG, pela passagem de conhecimento durante toda a graduação e pelos ensinamentos que levarei para a vida. Por fazer um curso que é meu sonho desde pequeno ser uma conquista ainda maior do que pensei que seria, mesmo que nesse tempo tenha sonhado tantas vezes.

À minha família, Julio, Josilene e Vinnicius, que não mediram esforços e não duvidaram nem por um momento dessa realização. Por todo o ambiente que me proveram para que eu sempre tivesse o foco nos meus estudos, além de todo incentivo e suporte que me foi oferecido, sem eles não seria capaz de trilhar um por cento desse caminho.

Ao professor Cláudio Campelo pela oportunidade, por ter sido um ótimo orientador e além disso foi essencial na construção deste documento.

Aos meus amigos da graduação, que estavam presentes comigo nos melhores momentos e também nos piores além disso, me forneceram as melhores dicas e aprendizados que eu poderia obter na

vida. Dentre eles Gustavo, Gildo, Gabi, Lourival, Guilherme, Túlio, Jeyvi e Sheila. Para eles, que passaram noites em claro ao meu lado, meu mais sincero obrigado. Além de todo o meu carinho, que os momentos tenham sido retribuídos e marcantes para todos da mesma forma.

Aos meus irmãos e irmãs de coração, que fiz graças ao meu caminho como estudante e que sempre estiveram torcendo por mim, Heriberto, Kleber, Lucas, Paloma, Madu, Juci, Bianca, Ruhama, Mariana e Helen. Por toda a coragem e sorrisos que foram incentivados em meio aos momentos mais turbulentos da graduação, agradeço de coração.

E novamente, agradeço a Heriberto por tudo. Sem a sua companhia e companheirismo, jamais teria concluído este sonho.

REFERÊNCIAS

- [1] Kristina Chodorow and Michael Dirolf. 2013. *MongoDB: The Definitive Guide*. O'Reilly Media.
- [2] Katspaugh. 2022. *Wavesurfer.js*. <https://github.com/katspaugh/wavesurfer.js>
- [3] Robert C. Martin. 2008. *Clean Code: A Handbook of Agile Software Craftsmanship*. Prentice Hall.
- [4] Jenifer Tidwell. 2023. *Designing Interfaces: Patterns for Effective Interaction Design*. O'Reilly Media.