

UNIVERSIDADE FEDERAL DE JUIZ DE FORA  
INSTITUTO DE CIÊNCIAS EXATAS  
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

# **Audioteca: Gerenciador de acervo de áudios**

**Gilson dos Reis Dias Fonseca**

JUIZ DE FORA  
JULHO, 2019

# Audioteca: Gerenciador de acervo de áudios

GILSON DOS REIS DIAS FONSECA

Universidade Federal de Juiz de Fora  
Instituto de Ciências Exatas  
Departamento de Ciência da Computação  
Bacharelado em Ciência da Computação

Orientador: Eduardo Barrére

JUIZ DE FORA  
JULHO, 2019

# AUDIOTECA: GERENCIADOR DE ACERVO DE ÁUDIOS

Gilson dos Reis Dias Fonseca

MONOGRAFIA SUBMETIDA AO CORPO DOCENTE DO INSTITUTO DE CIÊNCIAS EXATAS DA UNIVERSIDADE FEDERAL DE JUIZ DE FORA, COMO PARTE INTEGRANTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE BACHAREL EM CIÊNCIA DA COMPUTAÇÃO.

Aprovada por:

Eduardo Barrére  
Doutor em Engenharia de Sistemas e Computação

Daves Martins  
Mestre em Computação de Alto Desempenho

Tatiane Ornelas Matins Alves  
Mestre em Ciência da Computação

JUIZ DE FORA  
10 DE JULHO, 2019

*Aos meus amigos e irmãos.*

*Aos pais, pelo apoio e sustento.*

## Resumo

Nos dias de hoje é perceptível o uso de diversos tipos de arquivos como difusores de informação através da internet, visto que essa propicia um canal rápido de disseminação. Temos arquivos de imagem, vídeo, áudio e etc. Porém tais arquivos possuem muito mais informações do que se aparenta. O arquivo de áudio por exemplo, a voz (texto falado) em si é uma informação que podemos tratar à parte por meio de transcrição automática, sendo também possível utilizar essa transcrição para a recuperação de tal arquivo. Outra possibilidade que se apresenta com a internet, é a criação de ambientes colaborativos, onde pessoas possam interagir e compartilhar suas ideias, bem como documentos e arquivos. Embora existam ambiente colaborativos que permitam o compartilhamento e uso de arquivos de áudio, como o Soundcloud, que permite que seus usuários armazenem um arquivo de áudio e outras pessoas possam comentar em trechos deste áudio, não foi identificado um ambiente que oferecesse de maneira transparente a transcrição em texto destes áudios. Por este motivo, este trabalho propõe um sistema que integre todas essas funcionalidades em um só lugar, com o intuito de agregar maior valor a informação repassada.

**Palavras-chave:** Ambiente Colaborativo, Transcrição automática de fala, Recuperação de informação.

# Abstract

Nowadays, the use of different types of files as information disseminators through the internet is perceivable. It happens because the internet provides a rapid channel of information dissemination. We have files that allow the dissemination of image, video, audio etc. But such files have far more information than we can see at first glance. In the case of an audio file for example, the voice (spoken text) itself is an information that we can treat separately by using automatic transcription, and it is also possible to use this transcription for the retrieval of such a file. Another possibility that presents itself with the internet is the creation of collaborative environments, where people can interact and share their ideas, as well as documents and files. Although there is collaborative environments that allows the sharing and use of audio files, such as Soundcloud, which allows its users to store an audio file and others can comment on snippets of it, an environment that transparently offers the transcription in text of these audios was not identified. For this reason, this work proposes a system that integrates those features in only one place, with the purpose of adding greater value to the information passed on.

**Keywords:** Collaborative Environment, Automatic Speech Recognition, Information Retrieval.

## Agradecimentos

Primeiramente aos meus pais, Gilson e Luiza Helena, por tudo o que fizeram e fazem por mim, pela educação que me deram e todo apoio ao longo dessa jornada, sem vocês não chegaria a lugar algum.

A todos os meus irmãos e irmãs, por estarem sempre ao meu lado.

Aos meus amigos, com os quais eu compartilhei essa jornada e que fizeram dela algo mais especial.

Ao professor Eduardo Barrére pela orientação, e principalmente, pela paciência, sem a qual este trabalho não se realizaria.

Aos professores do Departamento de Ciência da Computação pelos seus ensinamentos e aos funcionários do curso, que durante esses anos, contribuíram de algum modo para o nosso enriquecimento pessoal e profissional.

*“Se tens tempo para pensar em uma bela morte, porque não vives belamente até o fim?”.*

*Gintoki Sakata (Gintama)*

# Conteúdo

<b>Lista de Figuras</b>	<b>7</b>
<b>Lista de Tabelas</b>	<b>8</b>
<b>1 Introdução</b>	<b>9</b>
1.1 Descrição do Problema . . . . .	9
1.2 Justificativa . . . . .	10
1.3 Objetivos . . . . .	10
1.4 Metodologia . . . . .	10
1.5 Estrutura do Trabalho . . . . .	10
<b>2 Revisão Bibliográfica</b>	<b>12</b>
2.1 Em Recuperação de Informação . . . . .	12
2.2 Automatic Topic Segmentation for Video Lectures Using Low and High-Level Audio Features . . . . .	13
2.3 Projeto Audioteca: Ensino de Leitura Mediado Pelo Uso do Audiolivro . .	16
2.4 Audioteca Virtual de Letras: tecnologia para inclusão . . . . .	17
2.5 CollabEduc: Uma Ferramenta de Colaboração em Pequenos Grupos para Plataformas de EaD . . . . .	18
2.6 Conclusão . . . . .	21
<b>3 Desenvolvimento</b>	<b>22</b>
3.1 Arquitetura do sistema . . . . .	22
3.2 Transcrição . . . . .	23
3.3 Comentários . . . . .	25
3.4 Busca . . . . .	26
3.5 Apresentação do sistema . . . . .	27
3.5.1 Dashboard . . . . .	28
3.5.2 Meus áudios . . . . .	29
3.5.3 Projetos . . . . .	31
3.5.4 Usuários . . . . .	33
3.5.5 Busca . . . . .	34
<b>4 Experimentos</b>	<b>35</b>
4.1 Discussão . . . . .	38
<b>5 Considerações Finais</b>	<b>40</b>
<b>Bibliografia</b>	<b>41</b>

## Lista de Figuras

3.1	Modelo do Banco de Dados . . . . .	23
3.2	Código de solicitação da transcrição . . . . .	24
3.3	Código de recuperação da transcrição . . . . .	25
3.4	Código do componente comentário . . . . .	26
3.5	Código de recuperação de trechos das transcrições . . . . .	27
3.6	Código para verificar se contém a informação buscada . . . . .	27
3.7	Arquitetura do sistema . . . . .	28
3.8	Tela Inicial do Sistema . . . . .	28
3.9	Tela do Dashboard com notificação de participação em projeto . . . . .	29
3.10	Tela Meus Áudios com lista de áudios . . . . .	29
3.11	Tela de inclusão de áudio . . . . .	30
3.12	Tela de anotações do áudio . . . . .	31
3.13	Tela de Projetos . . . . .	31
3.14	Tela de detalhes do projeto . . . . .	32
3.15	Tela de inclusão e edição do projeto . . . . .	33
3.16	Tela de usuários cadastrados no sistema . . . . .	33
3.17	Tela de Busca . . . . .	34

## Lista de Tabelas

4.1	Respostas ao Questionário . . . . .	38
-----	-------------------------------------	----

# 1 Introdução

A internet permite a rápida disseminação de informação entre seus usuários. Entre essas informações podemos destacar arquivos de áudio, que permitem que sons sejam transmitidos por esta via rápida de comunicação.

A voz (o texto falado), que potencialmente poderá estar contida em um arquivo de áudio, em si é uma informação que podemos tratar a parte. Uma maneira de obtermos esta informação em forma de texto é fazermos uso de um processo automático de transcrição. Há várias maneiras de fazer tal processo, mas, basicamente, o áudio é analisado e, identificado a voz do locutor, as ondas sonoras serão comparadas com um banco de dados de palavras cujas sinais de áudio são conhecidos e, apresentando similaridade, o transcritor gerará esta informação em forma de texto.

Outra possibilidade que se apresenta com a internet, é a criação de ambientes colaborativos virtuais. Tais ambientes possibilitam que, mesmo longe fisicamente, pessoas possam interagir e compartilhar recursos, como documentos, arquivos de áudio, imagens etc. Esses ambientes virtuais podem ser aplicados em diversos setores, como na educação (salas virtuais com material educativo compartilhado entre os usuários) e no corporativo (salas de reuniões virtuais com documentos e apresentações compartilhados).

## 1.1 Descrição do Problema

Embora a internet possibilite tamanha facilidade de troca de informações, ela também faz com que esse grande fluxo gere dificuldades na recuperação de arquivos relevantes, por falta de metadados que possam dar mais semântica a estes arquivos e facilitar a identificação dos mesmos.

## 1.2 Justificativa

A existência de metadados gerados a partir de comentários associados a trechos do áudio e de um texto transcrito junto ao áudio, aumentaria a quantidade de informações associadas à mídia, facilitando entre outras coisas a busca em um repositório de áudios.

## 1.3 Objetivos

O objetivo almejado neste trabalho é a criação de um sistema web de compartilhamento de arquivos de áudio em um ambiente colaborativo virtual, denominado Audioteca, onde se permita ao usuário fazer comentários em trechos do áudio e a transcrição automática destes áudios, gerando assim metadados para a identificação dos mesmos. Também é objetivo deste trabalho criar funções de busca que possibilitem a recuperação de tais arquivos de áudio.

## 1.4 Metodologia

Como passo inicial para a formulação do projeto, foi realizada uma revisão bibliográfica dos conceitos e trabalhos relacionados no intuito de identificar pontos relevantes para obtenção de bons resultados, abordada com maior profundidade no capítulo 2.

Após a definição dos conceitos e dos pontos mais relevante, é proposto um sistema, que combina essas técnicas de forma a englobar tais conceitos em uma única ferramenta. O capítulo 3 descreve como foi feita a implementação de tal sistema assim como apresenta a arquitetura do mesmo.

Com a implementação do sistema, foi realizado um experimento em ambiente prático de uso. Os dados experimentais terão sua coleta e análise explanadas no capítulo 4.

## 1.5 Estrutura do Trabalho

Este trabalho é estruturado em cinco capítulos, os quais estão divididos em Introdução, Revisão Bibliográfica, Desenvolvimento, Experimentos e Considerações Finais.

---

A Introdução apresenta o tema principal, com um breve contexto histórico, além de descrever os problemas a serem tratados, justificar a abordagem para tratá-los e estipular os objetivos.

O capítulo de revisão bibliográfica se divide em revisão dos estudos realizados na literatura, de modo a definir termos e conceitos utilizados, e trabalhos relacionados com o objetivo de se conhecer o que já foi desenvolvido bem como suportar a definição deste trabalho. Em seguida o capítulo Desenvolvimento trata o desenvolvimento, abordando as ferramentas utilizadas, tal como a apresentação do sistema explicando as principais funcionalidades do mesmo.

Já o capítulo de Experimentos apresenta a dinâmica e resultados do uso da Audioteca por um grupo de alunos, com intuito de avaliar os possíveis benefícios que o Audioteca pode acrescentar à educação. Por fim, o capítulo de Conclusões apresenta as considerações finais sobre todo o processo de desenvolvimento do trabalho.

## 2 Revisão Bibliográfica

Este capítulo apresenta trabalhos científicos que abordam os assuntos envolvidos neste trabalho. Será apresentado o uso de arquivos de áudio em diferentes áreas de ensino em prol de embasar a eficácia de sua utilização, assim como conceitos de estudos colaborativos, recuperação da informação e transcrição automática.

### 2.1 Em Recuperação de Informação

(CARDOSO, 2004)

Neste artigo é dito que a recuperação da informação estuda o armazenamento e recuperação automática de documentos, que são objetos de dados. O processo de recuperação consiste na geração de uma lista de documentos recuperados para responder a consulta formulada pelo usuário. Um sistema de recuperação da informação (SRI) classifica os documentos recuperados, de acordo com uma ordem de relevância. Avalia-se o SRI através da comparação das respostas geradas por este sistema e o conjunto ideal de respostas. Essa comparação gera dois índices de avaliação: precisão e revocação. Onde a precisão é a fração dos documentos já examinados que são relevantes e a revocação é a fração dos documentos relevantes observada dentre os documentos examinados.

São então apresentados modelos clássicos, utilizados no processo de recuperação da informação (booleano, vetorial e probabilístico) que apresentam estratégias de busca de documentos relevantes para uma consulta.

O modelo booleano, que dada uma consulta, recupera documentos que contêm os termos que satisfazem a expressão lógica da mesma. Para estabelecer uma ordem de relevância, índices são criados baseados na comparação entre a consulta e os documentos. Neste modelo, uma consulta é considerada como uma expressão booleana convencional formada com os conectivos lógicos AND, OR e NOT.

O modelo vetorial gera um vetor resultado através de um cálculo de similaridade que se baseia na proximidade entre termos dos vetores documento e consulta. A proximi-

dade, por sua vez, é calculada ao se atribuir pesos que especificam o tamanho e a direção de seu vetor de representação.

O modelo probabilístico descreve documentos considerando pesos binários que representam a presença ou ausência de termos, e o vetor resultado é gerado a partir da probabilidade de que um documento seja relevante para uma consulta. Este modelo se baseia no princípio probabilístico de ordenação, que acredita na hipótese de que a relevância de um documento para determinada consulta é independente de outros documentos.

Uma outra estratégia para a recuperação da informação apresentada é a realimentação de relevantes, onde o usuário formula sua busca inicial, são retornados os documentos condizentes a busca e então o usuário seleciona os de sua preferência. O sistema então seleciona termos pertencentes aos documentos e os utiliza para reformular a consulta.

Além das estratégias para a recuperação da informação anteriormente apresentadas pelo autor, são apresentadas técnicas que tem como objetivo melhorar a resposta à consulta. A passagem, que é uma forma de apresentar o documento retornado ao usuário em porções menores. A expansão de consultas por meio da busca de palavras semelhantes aos termos da consulta. A filtragem de informação, que compara documentos com tópicos de interesse do usuário e retorna os mais pertinentes. A categorização de documentos, que os classifica em categorias bem definidas. Por fim tem-se a visualização, que foca no desenvolvimento de interfaces simples e amigáveis.

Ao fim do trabalho foi concluído que o estudo da área de recuperação de informação é de grande utilidade para a comunidade de sistemas de informações em geral, e que de fato modelos para a recuperação precisa de informações passaram a ser de muito maior importância.

## **2.2 Automatic Topic Segmentation for Video Lectures Using Low and High-Level Audio Features**

(SOARES; BARRÉRE, 2018)

O artigo apresenta um método para segmentação de vídeo aulas em tópicos de

forma automática, combinando funcionalidades de alto nível de anotação de semântica e busca em uma base conhecida com baixo nível de funcionalidades de áudio, assim como: detecção de silêncio, transcrição de áudio e frequências fundamentais. De forma a obter uma segmentação de tópicos acurada.

O processo pode ser visto como um pipeline, onde a saída de um estágio é a entrada do próximo. O processo se inicia com o vídeo tendo seu áudio extraído, e posteriormente este áudio é dividido em pedaços de forma que estes pedaços não contenham silêncio e, assim, temos então uma sequência de pedaços de áudios completamente preenchidos com fala. Após isso, são extraídos destes pedaços vetores que remete às frequências fundamentais e a densidade de força espectral e assim os termos são anotados, e para cada termo  $am$  pertence  $A_i$ , onde  $m$  pertence  $[0,1,2,\dots,G_i-1]$  e  $G_i-1$  é o número de termos anotados em  $A_i$ , é feita uma busca na base conhecida para encontrar o mesmo significado do termo  $am$ . Após isso, a categoria dos termos, incluindo  $am$ , é pega da mesma forma. Ao fim deste processo, para cada termo  $am$  temos seu sinônimo e sua categoria. Então neste passo é construído um vetor de espaço e calculada uma matriz de afinidade  $M_a$ . No passo seguinte as matrizes de afinidade são linearmente combinadas de forma a gerar uma única matriz  $H$ , usada para agrupar os pedaços de áudio. Então, a partir de timestamps, os grupos de áudio são mapeados em tópicos nas vídeo aulas.

Para a obtenção dos conteúdos de fala em forma de texto dos pedaços de áudio o autor utiliza o processo de automatização do reconhecimento de fala (ASR). Tal processo, então, é definido como um processo que recebe um áudio  $S$  como entrada e retorna uma sequência de palavras de um vocabulário que tem maior probabilidade de ser o que foi dito em  $S$ . Para isso um sistema ASR geralmente segue uma arquitetura composta por uma interface de extração, um modelo acústico, um modelo léxico e modelo linguístico. A interface de extração é responsável por gerar amostra do sinal de áudio e obter vetores que representam esses sinais. O modelo acústico relaciona estes vetores com fenômenos da linguagem. O modelo léxico relaciona as sequências de fenômenos reconhecidos com uma sequência de palavras de um vocabulário. Por fim o modelo linguístico serve para melhorar os resultados, prevenindo que sentenças que não possuem gramática sejam formadas.

Anotações semânticas para identificar entidades no texto transcrito e conseguir

suas referências na base conhecida são feitas e posteriormente outra busca na base conhecida é feita visando encontrar palavras relacionadas, fazendo que com isso o conhecimento sobre o conteúdo do vídeo seja extrapolado.

Para a construção das matrizes de afinidade o autor se vê na necessidade da criação de três vetores. O primeiro representa os pedaços de áudio em termos de variações de frequências fundamentais e densidade de força espectral. O segundo representa os pedaços de áudio em termos de funcionalidade de transcrição. O terceiro representa as anotações de semântica. Assim são geradas três matrizes de afinidade que representam cada um dos vetores, após o cálculo de afinidade destas matrizes com os vetores é gerada a matriz final  $H$  que representa a afinidade de cada par de pedaço de áudio, considerando a ocorrência múltipla de todos os níveis de informação.

Por fim utilizando-se do algoritmo para agrupamento espectral são gerados grupos de pedaços de áudio que representam os tópicos de uma vídeo aula. Após a obtenção do melhor agrupamento e obtida uma lista com os pedaços de áudio, onde cada um deles possui um label de agrupamento adicionado. Assim, de maneira sequencial, é feita a adição de pedaços de áudio com mesmo label em um mesmo tópico até que um label diferente seja avaliado, neste caso um novo tópico será gerado e o processo se reinicia.

Ao final do processo, utilizando timestamps, que indicam o tempo no vídeo em que os pedaços de áudio começam, é possível segmentar o vídeo educacional com os tópicos gerados pelo agrupamento gerado. Assim pode-se relacionar os limites de tópicos de áudio com os limites dos tópicos de vídeo com sucesso.

Foram realizado experimentos em vídeo aulas em português para evidenciar que com o método apresentado é possível obter segmentação de tópicos acuradas na maioria dos casos.

O artigo se encerra concluindo que ao fim da realização dos experimentos foi possível averiguar que a combinação de técnicas de baixo e alto níveis, melhoraram a segmentação do vídeos, confirmando assim que o método é capaz de segmentar os vídeos com sucesso.

## 2.3 Projeto Audioteca: Ensino de Leitura Mediado Pelo Uso do Audiolivro

(LIMA; MOURA, 2016)

O artigo apresenta o projeto Audioteca que foi aplicado a uma escola pública de ensino profissional do estado do Ceará, com, inicialmente, onze alunos do segundo ano do Ensino Médio e consistia em utilizar o audiolivro (texto em áudio) como mediador das aulas de leitura e como instrumento pedagógico a fim de trabalhar a desenvoltura oral dos alunos, desenvolver habilidades de interpretação textual e estimular a leitura.

O projeto foi iniciado em maio de 2012, os alunos se dividiram em quatro grupos baseados nas turmas em que estavam e ficou acordado que cada grupo leria uma obra literária do Romantismo ou do Realismo.

As obras foram escolhidas mediante propostas dos próprios alunos a fim de estimular o gosto pela leitura, e foi estipulado um prazo de um mês para que os alunos as lessem. Antes do final do prazo um dos alunos apresentou a professora uma gravação que ele mesmo havia feito, de trechos iniciais de *O Navio Negreiro*, de Castro Alves. O áudio foi produzido com fundo musical tenso, e o aluno realizou leitura dramática, com expressividade e clareza, surpreendendo a professora por sua autonomia e criatividade.

No decorrer das gravações ocorreram algumas dificuldades, desde o ruído que o ar-condicionado emitia até o fato de alguns alunos, por estarem atarefados com atividades escolares, necessitarem a alguns encontros previstos no cronograma. Por outro lado um dos grupos solicitou mais encontros devido à extensão da obra a qual ficaram encarregados, assim como sugeriu a participação de mais uma aluna ao projeto. Foi criado um blog onde resultados parciais do projeto eram divulgados a fim de que a comunidade escolar pudesse ter contato com o desenvolvimento e sugerisse melhorias.

Após três meses de atividades foi sugerido que o aluno deficiente visual da escola apreciasse as gravações e fizesse uma análise crítica do que estava sendo produzido até aquele momento.

Ao fim do processo estabelecido foram produzidos a quantidade de audiolivros estabelecida no início e ao longo do projeto, foi percebido uma leitura mais desenvolta

por parte dos alunos participantes, onde os mesmos conseguiram atentar para o uso da norma culta, demonstrando especial atenção ao ritmo e à entonação necessários a cada texto literário. O projeto conseguiu estimular, também, alunos que não participavam do mesmo a dar sua contribuição, gravando parte dos textos sugeridos bem como textos próprios.

Como conclusão, o artigo aponta que com o uso do audiolivro nas aulas de leitura, os alunos têm mostrado maior capacidade de concentração, reflexão e um aprendizado mais significativo.

## **2.4 Audioteca Virtual de Letras: tecnologia para inclusão**

(FONTANA; NUNES, 2005)

A Audioteca Virtual de Letras surgiu como projeto complementar à Biblioteca Virtual de Letras, no ano de 2004, e tem como objetivo fundamental, disponibilizar, através da internet, arquivos de áudio com leituras de trechos ou mesmos de obras completas da literatura de língua portuguesa, bem como da literatura universal, para que pessoas com deficiência visual possam ter acesso a tais obras.

O artigo trás como principal motivo para a concepção da Audioteca o fato de que os direitos básicos das pessoas com deficiência, no Brasil, estão muito longe de se tornarem uma realidade, principalmente se estas vivem afastadas dos grandes centros urbanos.

Foram avaliados métodos consagrados de facilitação do contato de deficientes visuais com a literatura, para assim, medir a relevância ou não do projeto. Dentre eles destacam-se livros em braile e programas de computador que sintetizam texto em texto falado, porém ambos os métodos foram tidos como insuficientes, pois apresentam alguns problemas desde alto custo até dificuldade na compreensão da voz sintetizada, assim o projeto avançou tentando sanar estes problemas.

Foi feita uma seleção para que os leitores, termo utilizado para referir-se às pessoas que realizam leituras para cegos, fossem os mais aptos para o trabalho. As gravações foram feitas em um pequeno estúdio que funciona na casa do professor coordenador do projeto

## 2.5 CollabEduc: Uma Ferramenta de Colaboração em Pequenos Grupos para Plataformas de EaD18

e a fim de remover qualquer ruído ou outras imperfeições apresentáveis nas gravações foi utilizado o software de edição de áudio Audacity, por ser simples tornando o trabalho ágil e eficiente. Ao fim dos trabalhos o arquivo é disponibilizado no site da Audioteca, cujo acesso é livre e grátis através da internet.

Em conclusão, o autor expressa que os resultados obtidos foram amplamente satisfatórios, onde se registrou um número expressivo de acessos ao site da Audioteca. Percebeu-se que não só deficientes visuais, mas também pessoas idosas, estudantes e professores, que utilizam o material em suas aulas, se beneficiaram da iniciativa.

O artigo termina informando que o curso de Licenciatura em Matemática a Distância estava dando os primeiros passos no sentido de aprovar um projeto semelhante de modo a disponibilizar em forma de áudio materiais didáticos da área.

## **2.5 CollabEduc: Uma Ferramenta de Colaboração em Pequenos Grupos para Plataformas de EaD**

(LIMA; SIEBRA, 2017)

Neste artigo é apresentado o CollabEduc que visa melhorar a aprendizagem dos alunos se apoiando na ideia de que os pares de alunos podem encontrar apoio um no outro na resolução de tarefas. Tentando, assim, sanar o problema de falta interatividade e estímulo aos alunos para com as Plataformas de Educação a Distância.

Para alcançar o objetivo almejado, o autor se utiliza das seguintes Teorias de Aprendizagem, no intuito de distinguir quais as condições essenciais e mais apropriadas para que a aprendizagem ocorra no indivíduo.

A Teoria de Piaget (PIAGET; MERLONE, 1976) que dispõe a ideia de que a inteligência do indivíduo é uma constante adaptação. A construção do conhecimento ocorre a partir de suas descobertas quando em contato com o mundo e com objetos.

Vygotsky (VYGOTSKY, 1979) que dispõe em sua teoria que o convívio e as interações sociais são uns dos fatores fundamentais para a aprendizagem. Ele postula ainda que o conhecimento é definido através do discernimento do que o indivíduo é capaz sem ajuda e o que ele pode fazer interagindo com o meio social.

## 2.5 CollabEduc: Uma Ferramenta de Colaboração em Pequenos Grupos para Plataformas de EaD19

Por fim, Siemens (SIEMENS, 2005) que defende a teoria de que os princípios e processos de aprendizagem são reflexos do contexto atual dos indivíduos.

Para a modelagem das estratégias de suporte à colaboração, foi utilizada a Grounded Theory ou Teoria Fundamentada em Dados, composta dos seguintes passos: etapa de comparação dos dados coletados (análise); identificação de palavra-chave a fim de sintetizar as informações (codificação); extração e agrupamento das similaridades destas informações fomentando conceitos (conceptualização), e agrupamento de conceitos semelhantes que podem emergir em algumas teorias (categorização). A base para coleta de dados se limitou à trabalhos da literatura que tratam das etapas de construção do conhecimento.

Na etapa de codificação, foram selecionadas palavras-chave para a formação de conceitos, como ponto de partida foram utilizados os textos das teorias de aprendizagem apresentadas. Na etapa de conceptualização, seguiu-se o modelo sistemático onde as palavras-chave são convertidas em conceitos.

O CollabEduc é formado por um player de vídeo, onde os membros do grupo de estudo terão acesso às videoaulas; Um quadro branco virtual este componente será responsável por apoiar as discussões entre os membros; E por fim um chat, onde os membros poderão interagir por meio de mensagens de texto.

Foram definidas estratégias de colaboração com o objetivo de perturbar a convergência natural dos seus utilizadores e proporcionar uma melhor aprendizagem. São elas a argumentação, que pode ser inserida neste contexto com o envio de questionamentos sobre o tema aos estudantes em pausas estratégicas no vídeo e, havendo divergência, os mesmos são convidados a arguir até chegar a um consenso; A explanação, onde os estudantes são separados em grupos e assistem diferentes vídeos e por fim, o explanam aos demais grupos; Por fim regulação mútua onde os estudantes, cada um em seu tempo pré-definido, contribuem para a solução de questões que se ativam ao fim da videoaula.

Após a definição dos parâmetros e métodos a serem utilizados na validação, foram definidos o público-alvo, os parâmetros comparativos, a especificação dos testes, a elaboração dos questionários e do relatório dos monitores. Foram recrutadas turmas do Ensino Superior com conhecimentos mínimos sobre algoritmos e programação. Ao todo

o experimento contou com 134 participantes divididos em 4 turmas. Cada turma foi dividida em dois grandes grupos (teste e controle) e, no grupo de teste, os participantes foram subdivididos em grupos menores de 4 a 5 participantes. Cada subgrupo recebeu uma lista de exercícios com 3 questões sobre um assunto pré-definido e foi definido que os participantes poderiam fazer uso de qualquer material de apoio. Para cada questão aplicada foi protocolada uma das estratégias de colaboração (argumentação, explanação ou regulação mútua). Após a aplicação de um questionário qualitativo sobre o uso das estratégias no experimento, foi aplicada uma avaliação com questões simples para os dois grandes grupos, sem o uso de material de apoio, com o intuito de verificar o nível da turma.

Terminada a fase de execução do experimento, iniciou-se a fase de consolidação dos resultados a partir da análise dos dados obtidos. Analisando-se os dados foi percebido a receptividade por parte dos estudantes ao conceito de grupos de estudos, bem como sobre a utilização de uma ferramenta para grupos de estudo online. Com dados em mãos notou-se que, na estratégia de argumentação, na primeira tentativa para resolver a questão, houve 53,32% de discordância em relação a resposta e que a maioria (89,6%) colaborou para chegar a um consenso.

Já na estratégia de Explanação, observou-se que 98,3% dos participantes interagiram e colaboraram na busca pela resolução da questão. Na estratégia de regulação mútua foi identificado que 75,43% dos participantes demonstraram atenção à atividade em algum momento dessa etapa do experimento. Ressalta-se que 64,18% não obedeceram ao protocolo e interromperam a contribuição do participante da vez a fim de auxiliá-lo.

Foi possível observar que 75% dos participantes sentiram-se estimulados a interagir com os colegas através do uso das estratégias, os participantes também concordaram fortemente que houve evolução no aprendizado ao utilizar as estratégias. Dentre as estratégias os participantes apontaram a colaboração por argumentação como a mais interessante e efetiva, pois o ato de primeiro resolver a questão sozinho e depois confrontar os resultados com o grupo traz a possibilidade de conhecer novas visões sobre a resolução de um mesmo problema ampliando o campo de conhecimento do estudante.

Através do experimento com os participantes, o autor conclui que as estratégias de

---

colaboração podem de fato estimular os alunos e auxiliar o ensino a distância no combate a evasão acadêmica.

## 2.6 Conclusão

Com a leitura dos artigos foi observado que o uso de transcrição, pode de fato auxiliar a dar mais significado à arquivos de áudio, desde a facilitar e criar oportunidades para o consumo do arquivo até um possível uso na indexação do arquivo para que posteriormente possa ser feita a recuperação do mesmo. Observou-se, também, que o uso de ambientes colaborativos estimulam os alunos de modo que os levam a interagir mais com o sistema e assim contribuir no entendimento do assunto abordado, com isso o potencial da junção dessas funcionalidades, no que se diz respeito a maior interação, maior entendimento e seu potencial uso para a educação, se torna um atrativo alto de forma a não poder ser ignorado.

## 3 Desenvolvimento

Este capítulo tem por objetivo apresentar o Audioteca. Um sistema web que permite usuários armazenarem áudios educacionais, criarem projetos com estes e outros áudios de conteúdos relacionados. Estes projetos então podem ser consumidos por outros usuários. Por intermédio de comentários em tempo real, o sistema visa estimular a arguição de modo a complementar os estudos.

### 3.1 Arquitetura do sistema

O Audioteca foi desenvolvida como um sistema web para que fosse possível ser acessada em computadores e dispositivos móveis, por meio de um navegador, tornando-a ubíqua. A implementação do sistema é baseado na estruturação full-stack em Javascript e pode ser dividido nas seguintes camadas: A interface de interação com o usuário (front-end), o acesso ao banco de dados em um servidor (back-end) e por fim uma camada que garante a comunicação entre as duas anteriores.

Para o desenvolvimento do front-end utilizou-se a biblioteca React implementada em Javascript, desenvolvida pelo Facebook e Open Source. O React tem como finalidade a criação de interfaces de usuários, com ele é possível a criação de componentes, assim como a reutilização de componentes criados por terceiros. Foram usados componentes do framework Material-UI, que utiliza como base os princípios do Material Design, de forma a facilitar o desenvolvimento e utilizar conceitos de bom design de interfaces, garantindo assim que a mesma fosse amigável e simples. Também foram criados componentes para funções específicas e para complementar e padronizar as interfaces, o CSS também foi utilizado.

O back-end foi desenvolvido utilizando a ferramenta Node.js que permite que o Javascript seja executado no servidor. Para complementar o Node.js foi utilizado o framework express.js para gerenciar as solicitações de diferentes requisições, rotas e URLs.

Quanto ao banco de dados, foi utilizado, inicialmente, o Firebase Realtime Da-

tabase que é um banco de dados NoSQL que utiliza esquemas no formato JSON para armazenar dados. Neste banco de dados é armazenado informações dos usuários, dos projetos e dos áudios (como pode ser visto na figura 3.1), porém no decorrer do projeto, foi identificada a necessidade de armazenar, também, os arquivos de áudio, para isso foi escolhido o Google Cloud Storage.

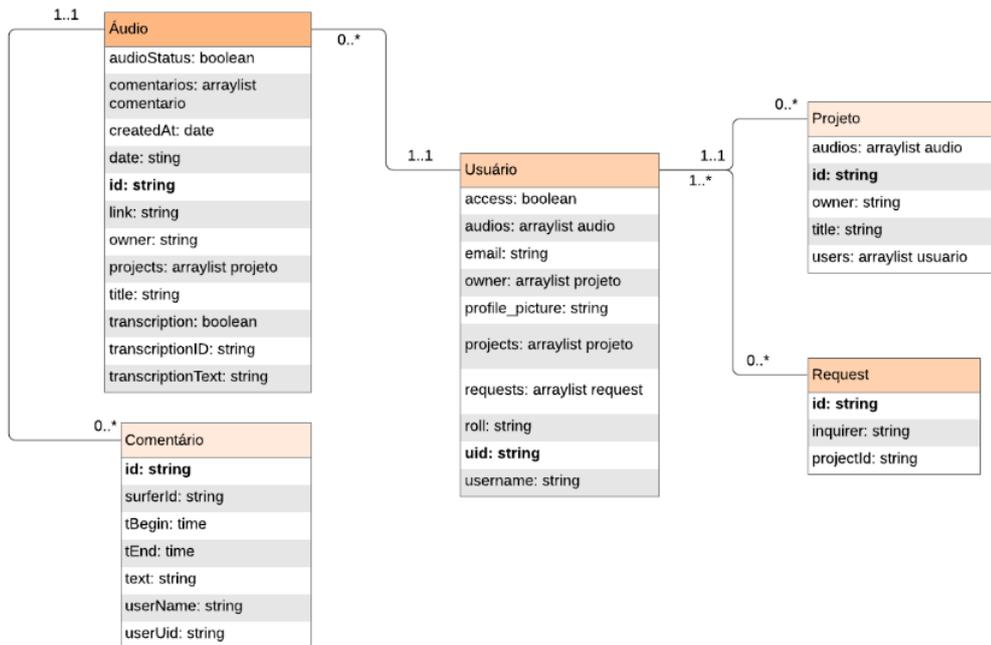


Figura 3.1: Modelo do Banco de Dados

## 3.2 Transcrição

Durante o desenvolvimento foi percebido a necessidade de integrar o sistema a um servidor para que fosse possível utilizar uma API de transcrição. Pelo pouco conhecimento do desenvolvedor nesta parte, foi necessário parar a produção para estudar quais ferramentas utilizar, como fazer e onde hospedar. Foi então escolhido montar um servidor em javascript com o Node.js e para hospedá-lo foi utilizada a plataforma do Google Cloud criando instâncias do Google App Engine.

A API de transcrição utilizada foi a Google Speech to Text. A decisão em se utilizar desta API em detrimento de outras foi devido ao conhecimento do desenvolvedor com o javascript e, pelo Firebase ser propriedade da Google o que facilitaria a integração das ferramentas.

Para que seja possível a transcrição nessa API é necessário que o arquivos esteja no formato WAV e desejado que a frequência de áudio seja de 16KHz. Foi utilizado, então, o ffmpeg que recebe o áudio e o formata para arquivo WAV. Por se tratar de uma função assíncrona foi necessário criar um fluxo que garantisse que a conversão fosse iniciada, apenas, ao fim do carregamento do áudio passado pelo usuário.

Ao fim da conversão do arquivo de áudio começa o processo de transcrição (caso aplicável), tal processo se dá em duas etapas. Na primeira (figura 3.2) etapa é criada uma configuração com informações a respeito ao formato do arquivo de áudio, bem como a língua falada. É também passado o link do repositório onde o áudio convertido se encontra, e com esses parâmetros é feita a solicitação à API da transcrição deste áudio. Neste momento é gerado uma ID (nome) que será utilizada na segunda etapa para a recuperação da transcrição.

Após ser cadastrado, o áudio que foi solicitada a transcrição, ficará inativo para edição ou acesso, neste período estará ocorrendo a segunda etapa (figura 3.3). Onde um listener verifica, por meio da ID gerada, se a transcrição foi finalizada, e assim que finalizada, a mesma é salva no banco de dados em tempo real do firebase e automaticamente o áudio fica disponível para acesso.

```
server > functions > JS gcslongspeechjs > <unknown> > asyncRecognize
1  var https = require('https');
2  const speech = require('google-cloud/speech').v1p1beta1;
3  const fs = require('fs');
4  var axios = require('axios');
5  var getToken = require('./jwt');
6
7  module.exports = async function asyncRecognize(path) {
8    const client = new speech.SpeechClient();
9    const gcsUri = `gs://audioteca/${path}`;
10   const encoding = 'LINEAR16';
11   const sampleRateHertz = 16000;
12   const languageCode = 'pt-BR';
13   const alternativeLanguageCodes = ['es-ES', 'en-US'];
14
15   const config = {
16     encoding: encoding,
17     sampleRateHertz: sampleRateHertz,
18     languageCode: languageCode,
19     alternativeLanguageCodes: alternativeLanguageCodes,
20     enableWordConfidence: true,
21     enableAutomaticPunctuation: true,
22   };
23   const audio = {
24     uri: gcsUri,
25   };
26
27   const request = {
28     config: config,
29     audio: audio,
30   };
31
32   const [operation] = await client.longRunningRecognize(request);
33   console.log('Speech gerado');
34   return JSON.stringify(operation.latestResponse.name);
35 }
```

Figura 3.2: Código de solicitação da transcrição

```
getTranscriptionAudio() {
  if(this.props.audiosList && this.state.isMounted) {
    Object.values(this.props.audiosList).map(async audio => {
      if(audio.transcription === true && (!audio.transcriptionText)) {
        var teste = await this.getTranscription(audio.transcriptionID);
        if(teste.done){
          if(teste.response.results !== undefined) {
            const transcription = teste.response.results
              .map(result => result.alternatives[0].transcript)
              .join('\n');
            firebase.database().ref(`/audios/${audio.id} /transcriptionText` )
              .set(transcription);
          }
          else {
            var msg = 'Não há nada para transcrever, tente com um áudio com fala em português brasileiro ou inglês americano';
            firebase.database().ref(`/audios/${audio.id} /transcriptionText` )
              .set(msg);
          }
        }
        else{
          setTimeout(() =>this.getTranscriptionAudio(), 60000);
        }
      }
    });
  }
}
```

Figura 3.3: Código de recuperação da transcrição

### 3.3 Comentários

Para promover um ambiente que permitisse a colaboração entre os usuários e, assim como a transcrição, fosse possível gerar metadados para a recuperação dos arquivos de áudio, foi desenvolvida a função de comentários.

Um comentário pode ser feito em anotações de áudio, e para fazê-lo o usuário deve selecionar um trecho do áudio e assim será permitido a inclusão de um comentário. O mesmo será salvo com o intervalo de tempo selecionado, e o nome do usuário que o fez. É possível editar e excluir o comentário feito, caso o usuário deseje. A imagem abaixo mostra um pouco como funciona a criação do componente comentário, é possível observar a função 'corrigeHora' que recebe um timestamp e o configura no formato HH:MM, é observável também que caso o usuário esteja editando o comentário aparecerá um textfield para preencher com o novo comentário, caso contrário aparecerá o comentário previamente gerado.



pelo usuário.

```
getTranscription(audios) {
  var list=[];
  Object.values(audios).map((audio, index)->{
    if(audio.transcriptionText && this.contains(audio.transcriptionText)) {
      var indice = audio.transcriptionText.toLowerCase().indexOf(this.state.searchField.toLowerCase());
      list[index] = ('...').concat(audio.transcriptionText.substring(indice, indice+200).concat('...'));
    }
    return null;
  });
  return list;
}
```

Figura 3.5: Código de recuperação de trechos das transcrições

```
contains(item) {
  if(item.toLowerCase().includes(this.state.searchField.toLowerCase()) && this.state.searchField !== '')
    return true;
  return false
}
```

Figura 3.6: Código para verificar se contém a informação buscada

## 3.5 Apresentação do sistema

Ao acessar o sistema pela primeira vez o usuário visualiza a tela de login (figura 3.8). Este login é gerenciado pelo Firebase Authentication e, no caso desta aplicação, a autenticação é feita utilizando um e-mail de domínio da Google. Para acessar qualquer rota no sistema, o usuário precisa estar autenticado, foi criado um componente que verifica se este é o caso, e caso contrário redireciona o usuário para a tela de login, garantindo assim maior segurança dos dados contidos no sistema. Um fluxograma do sistema pode ser observado na figura 3.7.

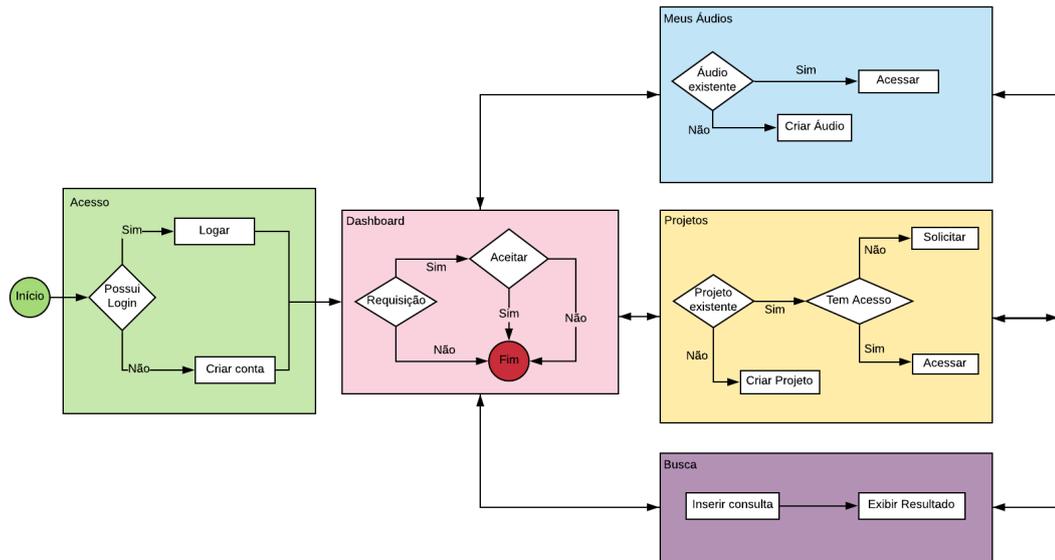


Figura 3.7: Arquitetura do sistema

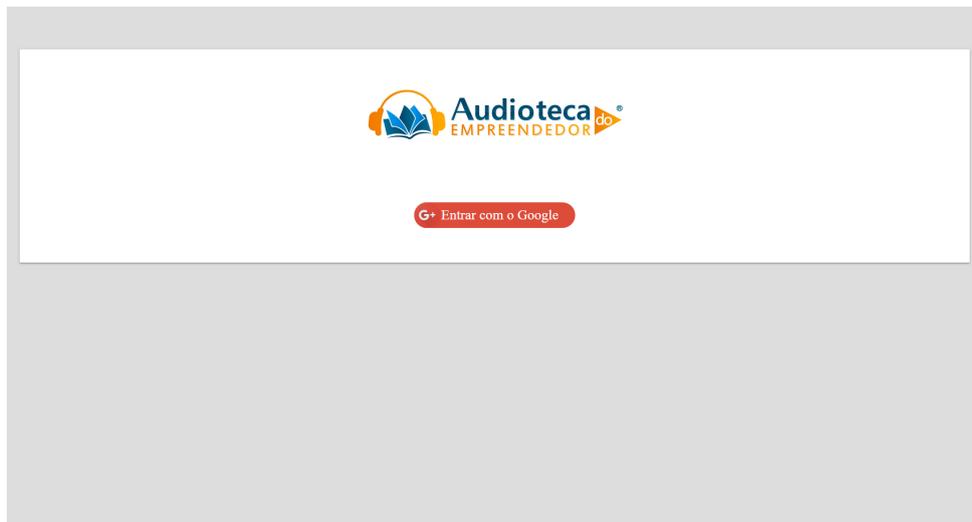


Figura 3.8: Tela Inicial do Sistema

### 3.5.1 Dashboard

Após se autenticar o usuário é redirecionado à tela "Dashboard" (figura 3.9), onde pode ver seus dados, como por exemplo, quantos áudios e projetos criou, quantos projetos participa e total de comentários feitos. A partir deste ponto o usuário pode navegar pela Appbar para as demais rotas.

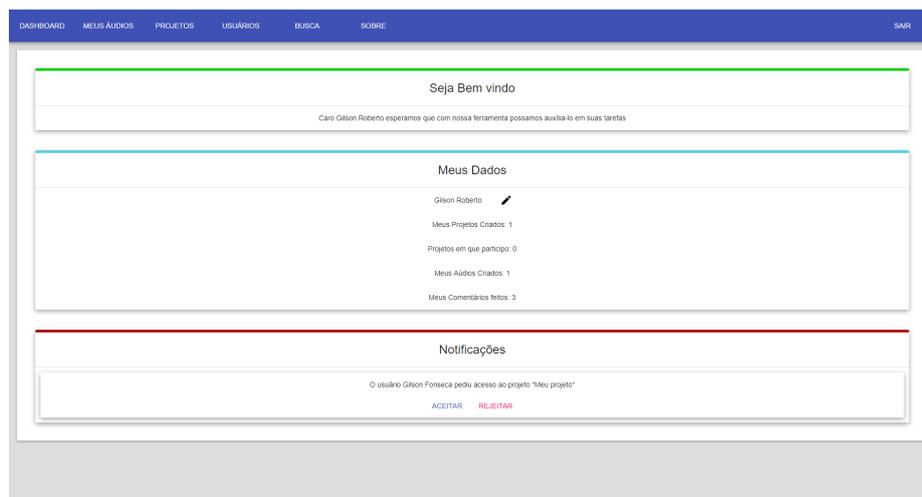


Figura 3.9: Tela do Dashboard com notificação de participação em projeto

### 3.5.2 Meus áudios

Em "Meus áudios" (figura 3.10) é mostrado uma lista com todos os áudios que o usuário criou, no caso do usuário ser um administrador do sistema aparecerão todos os áudios cadastrados. Neste menu é possível visualizar se o áudio é transcrito, se o mesmo foi encerrado e a quantidade de comentários feitos, assim como incluir um novo áudio, editar ou visualizar anotações de um áudio existente e removê-lo. Caso o áudio seja transcrito, porém a transcrição ainda não foi finalizada o áudio ficará indisponível para qualquer ação, e um listener verificará se a transcrição foi finalizada, em caso positivo o texto será adicionado ao áudio, caso contrário o áudio continuará indisponível.

Título	Transcrito	Finalizado	Comentários	Anotações	Editar	Excluir
PRISCILA	SIM	NÃO	0	👁️	✏️	🗑️
PRISCILA	SIM	NÃO	0	👁️	✏️	🗑️
PRISCILA RODRIGUES	SIM	NÃO	0	👁️	✏️	🗑️
PRISCILA RODRIGUES	SIM	NÃO	0	👁️	✏️	🗑️
UNAGI WAV	SIM	SIM	1	👁️	✏️	🗑️
KALIMBA WAV	NÃO	NÃO	1	👁️	✏️	🗑️
WILDLIFE (ONLINE-AUDIO-CONVERTER.COM) WAV	SIM	NÃO	0	👁️	✏️	🗑️
WILDLIFE (ONLINE-AUDIO-CONVERTER.COM) WAV	SIM	NÃO	0	👁️	✏️	🗑️
PROJETO 1 (ONLINE-AUDIO-CONVERTER.COM) WAV	SIM	SIM	2	👁️	✏️	🗑️
WATCHYVW-EKZLMSDQJQ	NÃO	NÃO	0	👁️	✏️	🗑️
BITUCA	NÃO	NÃO	0	👁️	✏️	🗑️
BITUCA	NÃO	NÃO	0	👁️	✏️	🗑️
SISTINE WAV	NÃO	NÃO	0	👁️	✏️	🗑️
VÍDEO A11 A JGGRS FFRICACIONAR WAV	SIM	NÃO	0	👁️	✏️	🗑️

Figura 3.10: Tela Meus Áudios com lista de áudios

Ao incluir ou editar um áudio (figura 3.11) é necessário informar o endereço do

áudio, ou selecionar um arquivo do computador, no segundo caso aparecerá uma barra de progressão informando a porcentagem de upload do arquivo selecionado. Após isso, ao clicar no botão confirmar, será possível adicionar um título ao áudio, porém este campo estará preenchido com uma sugestão baseando-se no conteúdo do endereço ou no nome do arquivo carregado, será possível selecionar também uma data, e se o usuário deseja que o áudio seja transcrito. A transcrição é feita a partir da API Speech to Text da Google, onde ela recebe um identificador e retorna o resultado da transcrição para este identificador. Ao solicitar a transcrição o identificador será adicionado ao objeto áudio.

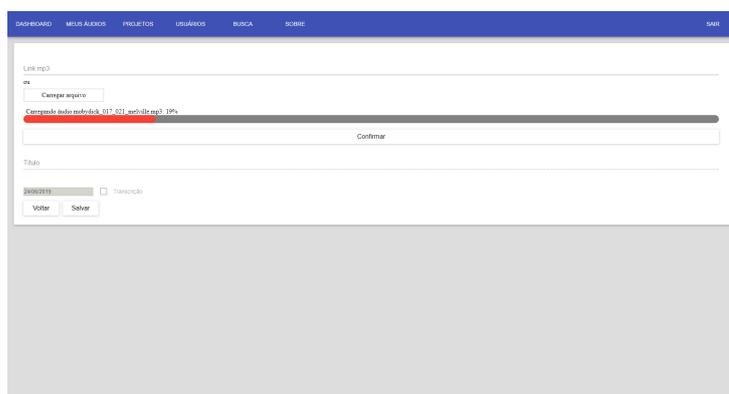


Figura 3.11: Tela de inclusão de áudio

Para remover um áudio este não pode estar sendo utilizado em nenhum projeto, caso isso ocorra uma mensagem será mostrada na tela, informando o motivo da não exclusão assim como informando a quem recorrer para auxílio, caso não ocorra, o áudio será excluído.

Ao visualizar anotações de um áudio (figura 3.12) o usuário terá acesso ao áudio em si, podendo ouvi-lo. Foi determinado que o player a ser utilizado para a reprodução do áudio seria o waversurfer.js, devido ao fato de que o mesmo possui métodos para selecionar trechos do áudio, devido a isso é possível criar comentários referentes a estes trechos. Nesta tela também é possível visualizar a transcrição, caso tenha sido solicitada, e os comentários feitos, caso haja algum, assim como editá-los e baixá-los em um arquivo CSV. Caso o usuário seja quem adicionou o áudio ao sistema, ele poderá encerrá-lo tornando indisponível adicionar ou editar comentários ou a transcrição.

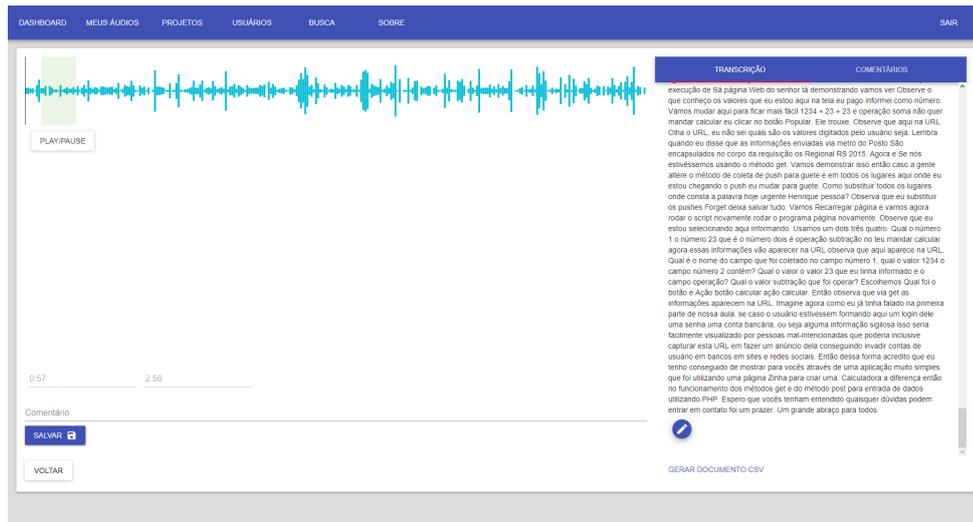


Figura 3.12: Tela de anotações do áudio

### 3.5.3 Projetos

Em "Projetos" (figura 3.13) o usuário visualiza uma lista com todos os projetos criados no sistema, assim como quem o criou e quantos participantes possui e pode efetuar as ações de incluir um novo projeto, ver detalhes de um projeto, solicitar a participação em um, ou excluir um projeto existente, caso seja o administrador deste projeto.

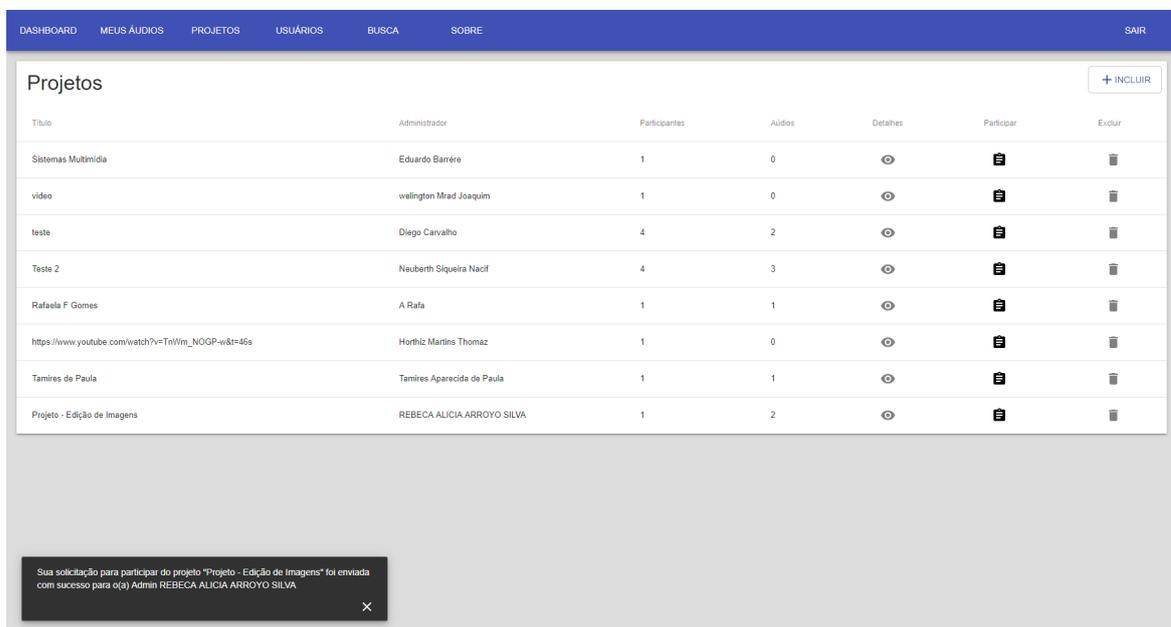


Figura 3.13: Tela de Projetos

Clicando em detalhes o usuário poderá visualizar os áudios e os participantes do projeto. E nos áudios poderá visualizar os detalhes do áudio, poderá também editar o projeto (conforme a figura 3.14).

Ao clicar em incluir ou em editar, o usuário será redirecionado para uma nova tela (conforme figura 3.15), onde, para finalizar o processo, deverá informar o nome do projeto, e por padrão já estará incluso no mesmo. O usuário pode também adicionar outros participantes que estejam cadastrados no sistema e áudios ao projeto, mas ambos os campos não são obrigatórios para o cadastro ou alteração do projeto.

O ícone de participar de um projeto, cria uma solicitação que é enviada ao administrador do projeto (figura 3.9), informando o interesse do usuário em participar do mesmo, caso o administrador aceite, o ícone ficará indisponível e o ícone de detalhes ficará disponível. Ao solicitar participação, uma mensagem aparecerá informando que o pedido foi enviado para o administrador (figura 3.13), para que o usuário receba um retorno de sua ação.

O último ícone a ser discutido é o de excluir um projeto. Não há, para a exclusão de um projeto, nenhuma limitação, sendo assim ao clicar no ícone, o projeto será excluído.

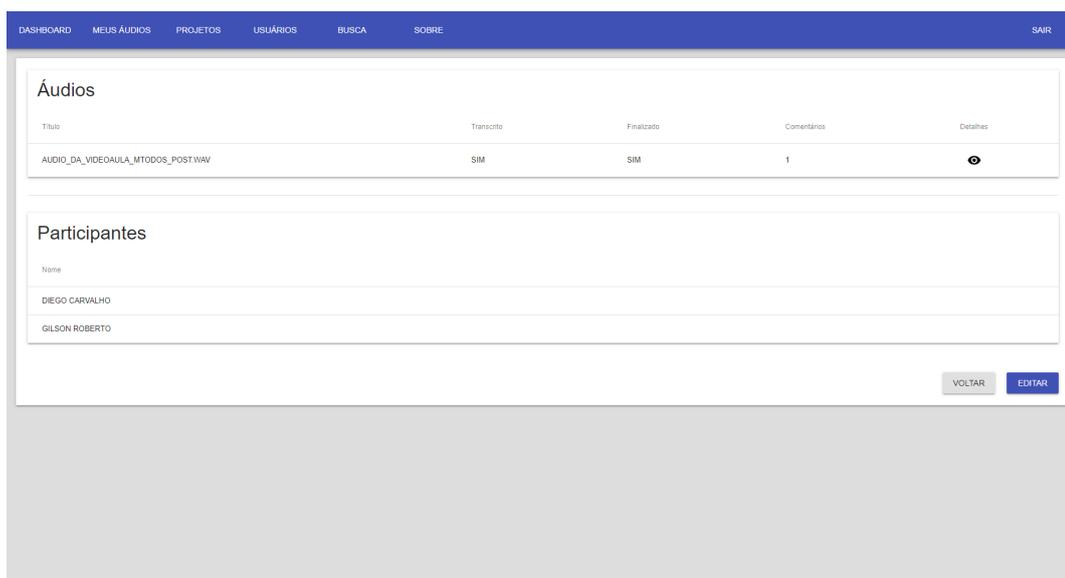


Figura 3.14: Tela de detalhes do projeto

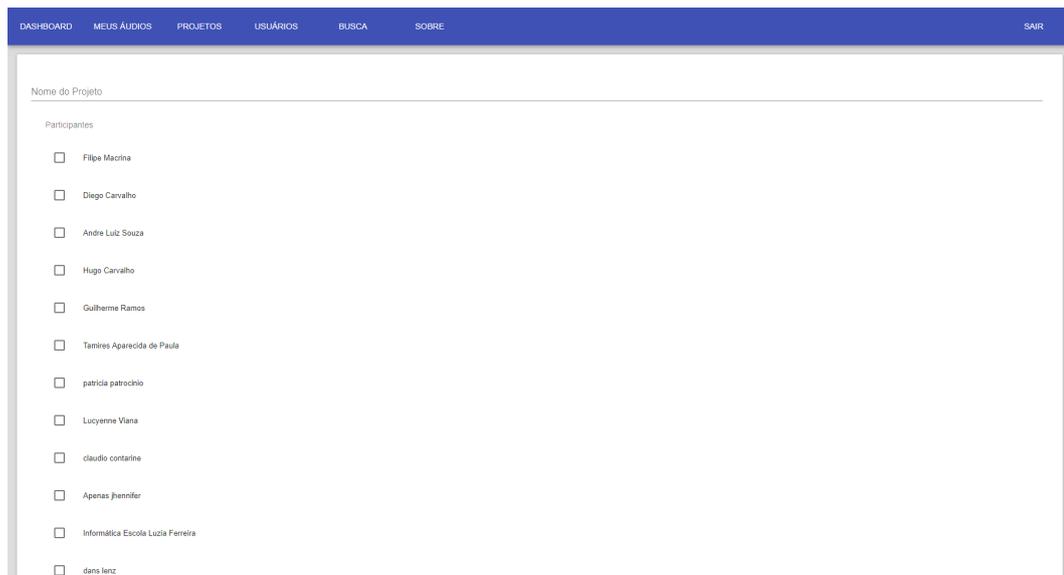


Figura 3.15: Tela de inclusão e edição do projeto

### 3.5.4 Usuários

Esta tela aparecerá apenas para os administradores, e possui uma lista com todos os usuários cadastrados no sistema assim como dois ícones, onde um serve para promover o usuário selecionado ao papel de administrador e o outro para rebaixa-lo para o papel de usuário.

Nome	Acesso	Tipo	Promover	Rebaixar
FILIFE MACRINA	SIM	User	↑	↓
DIEGO CARVALHO	SIM	User	↑	↓
ANDRE LUIZ SOUZA	SIM	User	↑	↓
HUGO CARVALHO	SIM	User	↑	↓
GUILHERME RAMOS	SIM	User	↑	↓
TAMIRES APARECIDA DE PAULA	SIM	User	↑	↓
PATRICIA PATROCINIO	SIM	User	↑	↓
LUCYENNE VIANA	SIM	User	↑	↓
CLAUDIO CONTARINE	SIM	User	↑	↓
APENAS JHENNIFER	SIM	User	↑	↓
INFORMÁTICA ESCOLA LUZIA FERREIRA	SIM	User	↑	↓
DANS LENZ	SIM	User	↑	↓
NEUBERTH SIQUEIRA NACIF	SIM	User	↑	↓
GILSON RORFRTO	SIM	Admin	↑	↓

Figura 3.16: Tela de usuários cadastrados no sistema

### 3.5.5 Busca

Dada uma consulta feita pelo usuário, um conjunto de projetos será retornado. É possível consultar por nome do áudio, por usuário, por trecho da transcrição, por comentários e por projeto.

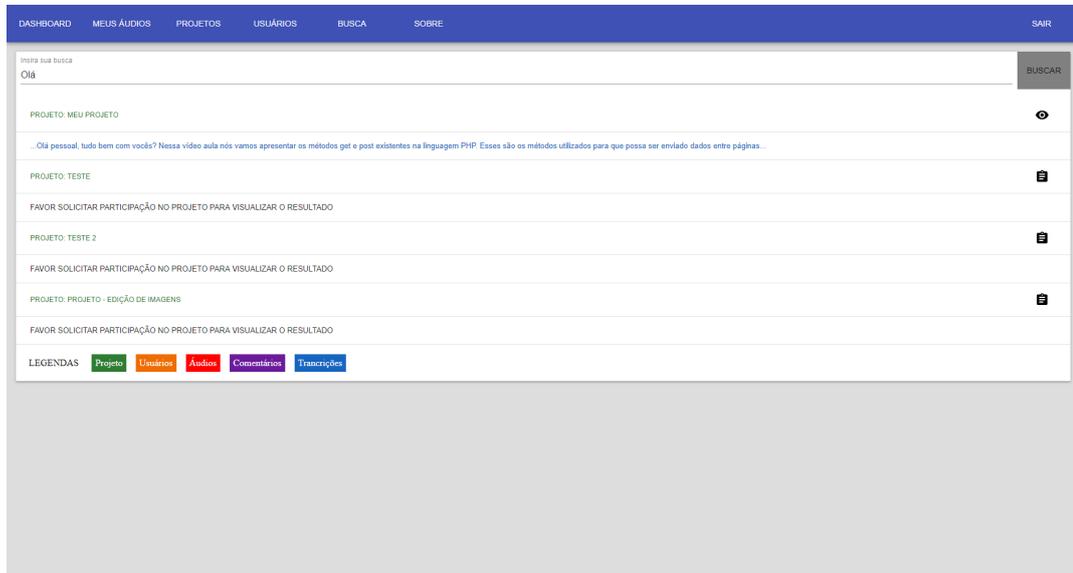


Figura 3.17: Tela de Busca

## 4 Experimentos

Para validar o sistema criado, foi realizado um experimento com uma turma de Sistemas de Multimídia do ensino a distância (EADDCC029-2019.1-A) da Universidade Federal de Juiz de Fora, onde um total de 43 alunos concordaram em participar.

Em um primeiro momento, tutoriais em formato de vídeo foram criados e distribuídos para os alunos participantes, estes tutoriais demonstram todas as funcionalidades do sistema e foram concebidos no intuito de familiarizar o aluno com o mesmo, sendo assim, foi solicitado aos participantes que antes de iniciarem o experimento visualizassem tais vídeos. Após se familiarizarem com o sistema, foi solicitado aos participantes a se cadastrarem para que lhes fosse permitido utilizá-lo.

O experimento, realizado durante o mês de Junho de 2019, consistiu na utilização do sistema por parte dos alunos. Estes fizeram uso da transcrição e geraram comentários em áudios, de forma a desenvolver atividades referentes ao seu curso.

Durante esse processo foi identificado que áudios advindos do Google Drive gravavam erro. Ao se verificar foi observado que a estrutura do link difere do padrão criado para a extração do áudio por links. Para este caso um novo padrão foi criado, porém para que fosse possível que o áudio fosse extraído era necessário a autenticação do usuário, por motivos de segurança do Google Drive e a inclusão de uma nova API. Para não aumentar a complexidade do sistema, foi então solicitado que o áudio fosse descarregado de modo manual pelo usuário em seu computador e pelo botão "Carregar Arquivos" fosse carregado ao sistema.

Durantes os experimentos também, foi verificado que requisições feitas ao Google App Engine que extrapolassem trinta segundos eram descartadas, sendo assim, arquivos com um tamanho superior a 20Mb não estavam podendo ser carregados para o sistema, para a resolução deste problema, foi decidido que inicialmente os arquivos devem ser carregados no banco de dados de armazenamento, pois o mesmo carrega arquivos de até 5Gb. Assim que forem carregados no banco de dados é feita a requisição ao servidor, este por sua vez carrega o arquivo do banco de dados o converte para o formato desejado, ao

fim do processo o arquivo original é excluído. Ao final do experimento, os alunos foram convidados a responder um questionário com sete questões referentes à sua experiência com o sistema.

A partir das respostas obtidas, pode-se observar um retorno positivo quanto a junção do uso da transcrição e comentários aos áudios, onde se nota um acordo de 95,35% por parte dos alunos quanto ao fato de que ambas são importantes e se completam.

É identificado, também, que os alunos concordam sobre o aspecto colaborativo do sistema como sendo uma ampliação de seu uso, onde podemos observar uma concordância de 90,70% e confirmar ao observarmos que os alunos veem tal aspecto como contribuinte quanto ao cenário educacional visto que a maioria (88,37%) acredita que o ambiente colaborativo criado possa contribuir no entendimento de áudios educacionais e ser uma boa ferramenta na gestão de podcasts (áudios) sobre a disciplina.

Quanto a transcrição automática é observado que os alunos tiveram como útil tal processo (62,79%), porém é também observado que o fato de poder editá-la para corrigir possíveis erros encontrados é bem recebido (27,91%), fazendo com que a aceitação de tal funcionalidade seja maior.

O fato de se poder recuperar arquivos de áudio utilizando determinados termos encontrados nos comentários e na transcrição faz com que o Audioteca seja avaliado, pela maioria dos alunos (86,05%), como útil para acervos de áudio.

Quanto a interface pode-se notar que 41,86% dos alunos a achou de fácil utilização já outros 27,91% tiveram uma dificuldade inicial, porém com a primeira utilização esta dificuldade foi sanada. Muitos (37,21%) acreditam que não houveram dificuldades no uso da ferramenta.

A tabela 4.1 traz as questões, assim como o resultado obtido pela avaliação dos alunos:

Questão 1 - No Audioteca o áudio pode ser anotado de forma automática e com comentários manuais. Sobre essas duas opções:	
Ambas são importantes e se complementam	95,35%
Somente a transcrição seria importante	4,65%
Somente a inserção de comentários seria importante	0,00%
A ferramenta não apresenta qualquer utilidade	0,00%
Questão 2 - O Audioteca permite termos a transcrição do áudio e também a geração de comentários a cerca de um determinando trecho desse áudio. Sempre associados a um intervalo temporal. Desta forma é possível realizar buscar por áudios que possuem um determinado termo (comentário ou na transcrição). Sendo assim, você considera o Audioteca:	
Útil para acervos de áudio	86,05%
De utilidade não bem clara para mim	13,95%
Sem qualquer aplicação clara para mim	0,00%
Questão 3 - O Audioteca, através do recurso de projetos (várias pessoas podendo interagir com diversos áudios), apresenta uma oportunidade de colaboração, pois mais de um usuário, cadastrado no mesmo projeto, pode realizar comentários em um mesmo áudio. Esse aspecto do Audioteca:	
Amplia o uso da ferramenta	90,70%
Não agrega valor em qualquer hipótese	6,97%
Por não ser útil em diversos casos, poderia não ter sido implementada	2,33%
Questão 4 - Uma dos usos do Audioteca pode ser em acervos de áudio educacionais. Assim, um professor pode criar um projeto, inserir diversos áudios e os alunos podem, de forma colaborativa (cada um faz os comentários que achar pertinente) contribuir no entendimento dos áudios. Neste cenário, o Audioteca utilizada para a educação pode:	
Ser uma boa ferramenta na gestão de podcasts (áudios) sobre a disciplina	88,37%
Não apresenta nada de novo no processo ensino-aprendizagem	9,30%
Fazer com que o professor perca o controle sobre o acervo de um determinado tipo de mídia disponibilizada.	2,33%

Questão 5 - O Audioteca faz uso da transcrição automática do áudio, sempre relacionada com um determinado tempo da áudio. Tal aspecto é:	
Útil para o processo de anotação do áudio	62,79%
Poderia ser interessante, mas a transcrição é muito ruim	6,98%
Mesmo a transcrição não sendo de ótima qualidade, o fato de poder editar cada transcrição, torna o processo útil	27,91%
Não agrega valor à ferramenta	2,32%
Questão 6 - Qual o grau de dificuldade em utilizar o sistema ?	
Nenhuma dificuldade	37,21%
Pouca dificuldade com a interface	27,91%
Pouca dificuldade com a manipulação do áudio	25,58%
Muita dificuldade com a interface	2,32%
Muita dificuldade com a manipulação do áudio	6,98%
Questão 7 - A forma para inserir comentários no áudio se mostrou:	
Fácil de ser utilizada	41,86%
Interface visual de marcação do áudio complexa	4,65%
Relação um pouco complicada entre texto a ser inserido e marcação visual	13,95%
Pouco intuitiva num primeiro momento	11,63%
Fácil de ser utilizada após a primeira vez	27,91%

Tabela 4.1: Respostas ao Questionário

## 4.1 Discussão

Em geral o sistema foi capaz de alcançar os objetivos deste trabalho, pois o mesmo auxiliou com sucesso as atividades propostas. Deve-se notar porém que, no quesito de manipulação de áudio e interação com a interface, houveram algumas dificuldades que somadas chegaram a 53,49% de respostas por parte dos alunos, permitindo concluir que

---

a interface apresenta falha quanto a inclusão e detalhes dos áudios e é um candidato a modificações.

## 5 Considerações Finais

Este trabalho, objetivou a construção de um sistema que permitisse ao usuário armazenar áudios, transcrevê-los e, recuperá-los por meio de busca, em um ambiente colaborativo. Foi então, desenvolvido o Audioteca, um sistema que unifica tudo o que se objetivou neste trabalho.

Com o término dos experimentos feitos com a turma de ensino a distância de sistemas multimídias foi possível verificar o correto funcionamento do sistema, pois os alunos foram capazes de utilizar todas as funcionalidades do mesmo. Atingindo assim o objetivo proposto de criar um sistema que contemplasse a transcrição automática, que permitisse a participação de seus usuários por meio de comentários e, a partir dos metadados gerados por essas duas ferramentas, que pudesse recuperar dados referentes a um áudio.

A partir do questionário foi possível identificar algumas dificuldades por parte dos alunos na utilização do sistema, como trabalhos futuros, pretende-se a melhoria das interfaces de interação com os áudios, tornando-as mais amigáveis. A melhoria das funções de busca, tornando-as mais inteligentes, também é almejada em trabalhos futuros.

## Bibliografia

- CARDOSO, O. N. P. Recuperação de informação. *INFOCOMP*, v. 2, n. 1, p. 33–38, 2004.
- FONTANA, M. V. L.; NUNES, E. V. Audioteca virtual de letras: tecnologia para inclusão. *RENOTE*, v. 3, n. 2, 2005.
- LIMA, E.; SIEBRA, C. Collabeduc: Uma ferramenta de colaboração em pequenos grupos para plataformas de aprendizagem a distância. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2017. v. 28, n. 1, p. 1707.
- LIMA, I. M.; MOURA, A. C. C. Projeto audioteca: Ensino de leitura mediado pelo uso do audiolivro. *Entrelinhas*, v. 10, n. 1, p. 20–33, 2016.
- PIAGET, J.; MERLONE, M. *A equilibração das estruturas cognitivas: problema central do desenvolvimento*. Zahar, 1976. (Ciências da educação). Disponível em: <https://books.google.com.br/books?id=TPActwAACAAJ>.
- SIEMENS, G. Connectivism: A learning theory for the digital age. In: *International Journal of Instructional Technology and Distance Learning*. [S.l.: s.n.], 2005.
- SOARES, E. R.; BARRÉRE, E. Automatic topic segmentation for video lectures using low and high-level audio features. In: ACM. *Proceedings of the 24th Brazilian Symposium on Multimedia and the Web*. [S.l.], 2018. p. 189–196.
- VYGOTSKY, L. S. *pensamento e linguagem*. Antídoto, 1979. Disponível em: <http://www.ebooksbrasil.org/adobeebook/vigo.pdf>.