

**CENTRO PAULA SOUZA
FACULDADE DE TECNOLOGIA DE FRANCA
“Dr. THOMAZ NOVELINO”**

TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE SISTEMAS

**GABRIEL MAIKON DA SILVA
GABRIEL MALTA CARDOSO**

**CIDA - CONVERSOR INSTANTÂNEO PARA DEFICIENTES
AUDITIVOS**

**FRANCA/SP
ANO 2023**

GABRIEL MAIKON DA SILVA
GABRIEL MALTA CARDOSO

CIDA - CONVERSOR INSTANTÂNEO PARA DEFICIENTES
AUDITIVOS

Trabalho de Graduação apresentado à Faculdade de Tecnologia de Franca - "Dr. Thomaz Novelino", como parte dos requisitos obrigatórios para obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas.

Orientadora: Profa. Dra. Jaqueline Brigladori
Pugliesi

FRANCA/SP
ANO 2023

GABRIEL MAIKON DA SILVA
GABRIEL MALTA CARDOSO

CIDA - CONVERSOR INSTANTÂNEO PARA DEFICIENTES
AUDITIVOS

Trabalho de Graduação apresentado à Faculdade de Tecnologia de Franca – “Dr. Thomaz Novelino”, como parte dos requisitos obrigatórios para obtenção do título de Tecnólogo em Análise e Desenvolvimento de Sistemas.

Trabalho avaliado e aprovado pela seguinte Banca Examinadora:

Orientador(a) : _____
Nome..... : Profa. Dra. Jaqueline Brigladori Pugliesi
Instituição : Faculdade de Tecnologia de Franca – “Dr. Thomaz Novelino”

Examinador(a) 1 : _____
Nome..... : Prof. Me. Leonardo Henrique Raiz
Instituição : Faculdade de Tecnologia de Franca – “Dr. Thomaz Novelino”

Examinador(a) 2 : _____
Nome..... : Prof. Esp. Alexandre Gomes da Silva
Instituição : Faculdade de Tecnologia de Franca – “Dr. Thomaz Novelino”

Franca, 06 de junho de 2023.

AGRADECIMENTO

Primeiramente, gostaríamos de agradecer às nossas famílias e amigos que sempre estiveram presentes durante toda a jornada acadêmica, nos apoiando nos momentos mais difíceis e compartilhando conosco as alegrias das conquistas. Vocês foram fundamentais para nos manter motivados e confiantes, mesmo diante dos desafios.

Gostaríamos também de agradecer a nossa orientadora Jaqueline, que sempre esteve disposta a ajudar e aconselhar em todas as etapas do trabalho. Sua sabedoria e experiência foram inestimáveis para a conclusão deste projeto.

Por fim, gostaríamos de agradecer aos professores, Ana e Carlão, que foram essenciais para o nosso desenvolvimento acadêmico e profissional. Suas aulas e orientações nos proporcionaram as ferramentas necessárias para enfrentar os desafios do mercado de trabalho e nos tornaram profissionais mais preparados.

A todos vocês, o nosso mais sincero obrigado!

Dedicamos este trabalho a todos os nossos amigos e família, que estiveram ao nosso lado, mesmo diante de todas as adversidades. As experiências, conversas e histórias que trocamos foram nossas principais motivações.

"Não é sobre dar voz aos que não têm voz, é sobre mudar o mundo para que todos possam ser ouvidos."

Greta Thunberg

RESUMO

Visando preencher a carência que existe no setor de acessibilidade e inclusão social, este trabalho apresenta a ideia de desenvolvimento de um conversor capaz de realizar a tradução do áudio de *streamings* e vídeos e transmiti-lo para a Língua Brasileira de Sinais (LIBRAS) para que possa ser renderizado em um modelo de assistente virtual 3D com o objetivo de facilitar o entendimento e perpetuar a inclusão social e digital para pessoas com deficiência auditiva. Para o desenvolvimento do projeto, foram utilizadas técnicas de reconhecimento de voz, processamento de linguagem natural e Deep Learning para identificação do conteúdo e transcrição em texto. Além disso, foram levantados requisitos e funcionalidades que atendem às necessidades dos usuários, assim como modelado o fluxo de execução até atingir o resultado exibido ao cliente. O projeto apresenta-se como uma solução inovadora para a inclusão de pessoas surdas em diferentes contextos, oferecendo benefícios como a promoção da acessibilidade, autonomia e redução das barreiras comunicacionais no âmbito educacional, social e do entretenimento.

Palavras-chave: Acessibilidade. Assistente virtual. Deficientes auditivos. Extensão. Língua de sinais.

ABSTRACT

Aiming to fill the gap in the accessibility and social inclusion sector, this work presents the idea of developing a converter capable of translating audio from streamings and videos and transmitting it to Brazilian Sign Language (LIBRAS) so that it can be rendered in a 3D virtual assistant model with the aim of facilitating understanding and perpetuating social and digital inclusion for hearing-impaired people. For the development of the project, voice recognition, natural language processing, and deep learning techniques were used for content identification and transcription into text. In addition, requirements and functionalities that meet users' needs were identified, and the execution flow was modeled until the result was displayed to the client. The project presents itself as an innovative solution for the inclusion of the deaf in different contexts, offering benefits such as promoting accessibility, autonomy, and reducing communication barriers in the educational, social, and entertainment spheres.

Keywords: Accessibility. Extension. Hearing impaired. Sign language. Virtual Assistant.

LISTA DE FIGURAS

Figura 1 – Hand Talk	25
Figura 2 – Assistentes Virtuais	30
Figura 3 – Matriz SWOT	35
Figura 4 – BPMN	37
Figura 5 – Diagrama de Caso de Uso	46
Figura 6 – Python	51
Figura 7 – Speech-to-Text	52
Figura 8 – VLibras	53
Figura 9 – FFmpeg	54
Figura 10 – Visual Studio Code	55
Figura 11 – Miro	56
Figura 12 – GenMyModel	56
Figura 13 – Fluxo CIDA	57
Figura 14 – Tela Youtube	59
Figura 15 – Tela Netflix	59
Figura 16 – Tela Twitch	60
Figura 17 – Tela Menu	61
Figura 18 – Tela Notificação.....	61

LISTA DE QUADROS

Quadro 1 – Línguas de Sinais	18
Quadro 2 – Componentes 5W2H	35
Quadro 3 – 5W2H.....	36
Quadro 4 – Requisitos Funcionais.....	38
Quadro 5 – Requisitos Não Funcionais	41
Quadro 6 – Regras de Negócios	43
Quadro 7 – Matriz de Rastreabilidade – RF X RN.....	45
Quadro 8 – Casos de Uso	47
Quadro 9 – Matriz de Rastreabilidade – RF X UC.....	50

LISTA DE SIGLAS

5W2H – *What, Why, Who, Where, When, How, How much*

API – *Application Programming Interface*

ASL – *American Sign Language*

AVI – *Audio Video Interleaved*

BERT – *Bidirectional Encoder Representations from Transformers*

BPMN – *Business Process Model and Notation*

BSL – *British Sign Language*

CIDA – *Conversor Instantâneo para Deficientes Auditivos*

ESG – *Environment, Social & Governance*

FLAC – *Free Lossless Audio Codec*

GCP – *Google Cloud Platform*

IBM – *International Business Machines*

IoT – *Internet of Things*

LBI – *Lei Brasileira de Inclusão*

LGP – *Língua Gestual Portuguesa*

LIBRAS – *Língua Brasileira de Sinais*

LSA – *Lengua de Señas Argentina*

LSCH – *Lengua de Señas Chilena*

LSF – *Langue des Signes Française*

MKV – *Matroska Video*

MP3 – *MPEG-1/2 Audio Layer 3*

MP4 – *MPEG Layer 4*

ODS – *Objetivos para o Desenvolvimento Sustentável*

PLN – *Processamento de Linguagem Natural*

SLN – *Sign Language of Netherlands*

SRT – *SubRip Subtitle*

STT – *Speech To Text*

SWOT – *Strengths, Weaknesses, Opportunities, Threats*

UML – *Unified Modeling Language*

WAV – *Waveform Audio File Format*

WFD – *World Federation of the Deaf*

SUMÁRIO

1 INTRODUÇÃO	14
2 LINGUA DE SINAIS	17
3 ASSISTENTES VIRTUAIS	20
3.1 TECNOLOGIAS DE RECONHECIMENTO DE VOZ.....	20
3.1.1 Empresas	21
3.1.2 Saúde	22
3.1.3 Segurança	23
3.1.4 Atendimento em <i>call centers</i>	24
3.2 TRADUÇÃO DE TEXTO PARA LINGUAGEM DE SINAIS.....	24
3.2.1 Processamento de Linguagem Natural	26
3.2.2 Hand Talk.....	27
3.2.3 FLAC	28
3.3 DESENVOLVIMENTO DE ASSISTENTES VIRTUAIS.....	29
3.3.1 Modelos de linguagem	31
3.3.2 Redes Neurais.....	31
3.3.3 Aprendizado de Máquina.....	32
3.3.4 Processamento de fala em assistentes virtuais.....	32
4 LEVANTAMENTO DE REQUISITOS E FERRAMENTAS	34
4.1 MATRIZ SWOT	34
4.2 5W2H	35
4.3 BPMN.....	36
4.4 REQUISITOS FUNCIONAIS	38
4.5 REQUISITOS NÃO FUNCIONAIS	41
4.6 REGRAS DE NEGÓCIOS	43
4.7 MATRIZ DE RASTREABILIDADE (REQUISITOS FUNCIONAIS X REGRAS DE NEGÓCIO)	44
4.8 DIAGRAMA DE CASO DE USO	45
4.9 DOCUMENTAÇÃO DE CASO DE USO.....	46
4.10 MATRIZ DE RASTREABILIDADE (REQUISITOS FUNCIONAIS X CASOS DE USO)	49

5 PROJETO CIDA	51
5.1 FERRAMENTAS	51
5.1.1 Python	51
5.1.2 Speech-To-Text.....	52
5.1.3 VLibras	53
5.1.4 FFmpeg.....	54
5.1.5 Visual Studio Code.....	55
5.1.6 Miro	55
5.1.7 GenMyModel	56
5.2 FUNCIONALIDADES DO SISTEMA	57
5.3 Telas do sistema	58
5.4 BENEFÍCIOS.....	62
5.5 POSSÍVEIS LIMITAÇÕES E DESAFIOS	63
CONSIDERAÇÕES FINAIS	64
REFERÊNCIAS.....	66

1 INTRODUÇÃO

O uso da tecnologia na acessibilidade, principalmente no âmbito de deficientes auditivos, é uma área que tem uma importância social muito grande, afinal, 80% de um total de cerca de 500 milhões de pessoas que têm deficiências auditivas (10 milhões apenas no Brasil) é semianalfabeta ou analfabeta funcional, segundo a *World Federation of the Deaf* (WFD). Sendo assim, apesar das plataformas de conteúdos audiovisuais disponibilizarem legendas como um meio de suprir a falta de acessibilidade e tornar possível o consumo do conteúdo por usuários surdos, a partir dos dados citados anteriormente, e se refletirmos sobre a situação desta população, a acessibilidade por meio de legendas não é um meio eficaz, pois exclui a população surda que está em estado de analfabetismo, já que uma grande parcela de pessoas que já nasceram com a deficiência ou perderam a audição antes de serem alfabetizadas na língua portuguesa têm dificuldade em ler as legendas (UNITED NATIONS, 2003).

Visando preencher essa carência no setor de plataformas de conteúdos audiovisuais e facilitar o cotidiano dessas pessoas por meio de um recurso de acessibilidade que seja de fato eficaz e que abrange a maior parcela possível da população, foi proposta a ferramenta Conversor Instantâneo para Deficientes Auditivos (CIDA), solução tecnológica capaz de, em tempo real, capturar o áudio falado em *streams*, converter para um formato de texto e enviar os dados resultantes para uma *Application Programming Interface* (API) que as utilize para a reprodução da linguagem de sinais por meio de um assistente virtual.

Foi usado como referência e modelo de estudo, o projeto desenvolvido por Andrade (2019) que utilizou técnicas, ferramentas e bibliotecas semelhantes às citadas neste trabalho e que, devido as amostras divulgadas em sua plataforma apresentarem resultados satisfatórios, optou-se por sugerir a sua utilização para o desenvolvimento da solução proposta neste artigo. Como concluído pelo autor, embora o resultado tenha se divergido da ideia inicial, eles foram bem promissores, com uma estatística de cerca de 80% de acertos nos testes realizados. Esses dados foram a justificativa principal para a implementação das técnicas citadas, priorizando eficiência e o bom funcionamento para os usuários.

Toda essa arquitetura e transformação para tornar o ambiente digital mais acessível por meio desta solução tem como principal objetivo quebrar a enorme barreira de comunicação entre as empresas e milhões de pessoas surdas que querem consumir *online*, já que se estima que apenas 1% dos sites no Brasil são considerados acessíveis atualmente (FORBES, 2021).

Um site acessível abre um novo canal de comunicação com milhões de pessoas, que se tornarão consumidores de produtos e plataformas de conteúdos audiovisuais (como o Youtube e a Twitch, por exemplo). Esse público movimentava cerca de R\$ 576,6 bilhões em renda própria por ano, como é citado no “Raio X da surdez no Brasil”, pesquisa organizada pelo Instituto Locomotiva. A partir disso, pode-se notar o quanto essa comunidade pode contribuir para o setor de *streaming*, um mercado que está em crescente ascensão e que se projeta que o aumento do consumo no país movimentará mais de R\$ 1 bilhão no setor em 2022. Mesmo que as maiores plataformas de *streaming* não informem seus números, os executivos do setor afirmam que os investimentos anuais por plataforma passam dos R\$ 100 milhões (ESTADÃO, 2019).

Por fim, vale mencionar que a acessibilidade nos sites é obrigatória para organizações públicas e privadas, conforme a Lei Brasileira de Inclusão (LBI):

Art. 63. É obrigatória a acessibilidade nos sítios da internet mantidos por empresas com sede ou representação comercial no País ou por órgãos de governo, para uso da pessoa com deficiência, garantindo-lhe acesso às informações disponíveis, conforme as melhores práticas e diretrizes de acessibilidade adotadas internacionalmente (BRASIL, 2015, p. 2).

Dito isso, conclui-se que ser acessível é contribuir diretamente para uma sociedade com mais diversidade e inclusão e com a ajuda de todos podemos tornar o ambiente digital um lugar melhor.

No Capítulo 2 serão abordados os principais detalhes que envolvem a linguagem de sinais, desde o seu significado até suas variações em outros idiomas. No Capítulo 3 tem-se a explicação e a abordagem do conceito de assistentes virtuais e toda a tecnologia envolvida em seu processo de criação e utilização, assim como um breve resumo de algumas áreas da Inteligência Artificial que a ferramenta CIDA faz uso. No Capítulo 4 são elucidadas as metodologias e os diagramas referenciados na engenharia de software, cujo objetivo é facilitar o entendimento do fluxo da aplicação, assim como seus principais requisitos. No Capítulo 5, referente ao projeto

CIDA, são mencionadas as principais ferramentas utilizadas, assim como exemplos de telas onde a aplicação proposta será executada. Por fim, tem-se as Considerações Finais, onde são citadas possíveis avanços e melhorias futuras na aplicação e os seus benefícios esperados.

2 LINGUA DE SINAIS

Neste capítulo serão abordadas algumas definições a respeito dos principais temas relacionados a proposta de criação da ferramenta de tradução instantânea. O objetivo é familiarizar e contextualizar sobre os conceitos mais importantes para entendimento do que está sendo proposto, não apenas em relação ao projeto em si, mas seu impacto no mercado.

Inicialmente, ao buscarmos entender o que é LIBRAS (Língua Brasileira de Sinais), é necessário antes, esclarecer o que é a língua de sinais.

Dado o contexto de criação e desenvolvimento das linguagens de sinais, tem-se a definição de Honora; Frizanco (2009, p. 41):

As línguas de sinais são naturais, pois surgiram do convívio entre as pessoas surdas. Elas podem ser comparadas à complexidade e à expressividade das línguas orais, pois pode ser passado qualquer conceito, concreto ou abstrato, emocional ou irracional [...]. Trata-se de línguas organizadas e não de simples junção de gestões. Por este motivo, por terem regras e serem totalmente estruturadas, são chamadas de LÍNGUAS. [...]. As línguas não são universais. Cada uma tem sua própria estrutura gramatical, sendo assim, como não temos uma única língua oral, também não temos apenas uma língua de sinais.

A Língua de Sinais é uma língua visual-espacial utilizada pela comunidade surda para se comunicar. É uma língua natural e completa, com gramática, sintaxe e estrutura próprias. Diferente das línguas orais-auditivas, que utilizam sons e palavras faladas, a Língua de Sinais utiliza sinais e expressões corporais.

Assim como existem várias línguas faladas no mundo, também existem várias línguas de sinais pelo mundo. Cada país tem sua própria língua de sinais, tal como temos nossa própria língua falada. Pode-se observar, no Quadro1, algumas línguas de sinais espalhadas pelo mundo.

Quadro 1 - Línguas de Sinais

País	Sigla	Descrição
	LIBRAS	Língua Brasileira de Sinais
	LGP	<i>Língua Gestual Portuguesa</i>
	SLN	<i>Sign Language of Netherlands</i>
	ASL	<i>American Sign Language</i>
	LSA	<i>Lengua de Señas Argentina</i>
	BSL	<i>British Sign Language</i>
	LSCH	<i>Lengua de Señas Chilena</i>
	LSF	<i>Langue des Signes Française</i>

Fonte: Adaptado de CRISTIANO, 2017, online

Assim como as línguas orais-auditivas, a Língua de Sinais é uma parte importante da identidade cultural da comunidade surda e deve ser respeitada e valorizada. Infelizmente, muitas vezes a língua de sinais é subestimada e considerada inferior às línguas orais-auditivas. No entanto, a Língua de Sinais é tão complexa e eficaz quanto qualquer outra língua, e deve ser reconhecida como tal (CRISTIANO, 2017).

É importante destacar que a Língua de Sinais não é uma linguagem universal, e não é compreendida por todas as pessoas surdas. Portanto, é essencial respeitar e valorizar a diversidade linguística da comunidade surda, e garantir que sejam

fornecidos recursos adequados para a comunicação e inclusão de pessoas surdas em todos os aspectos da vida social.

LIBRAS é a língua de sinais usada pela comunidade de surdos no Brasil e já foi reconhecida por lei. A lei que dispõe sobre a língua de sinais é a Lei nº 10.436, de 24 de abril de 2002. De acordo com o próprio termo, a Libras é utilizada somente no Brasil, assim como a Língua Portuguesa:

Parágrafo único. Entende-se como Língua Brasileira de Sinais - Libras a forma de comunicação e expressão, em que o sistema linguístico de natureza visual motora, com estrutura gramatical própria, constitui um sistema linguístico de transmissão de ideias e fatos, oriundos de comunidades de pessoas surdas do Brasil (BRASIL, 2002, p. 23).

Pelo exposto em Lei, deve-se pensar na LIBRAS como um idioma de mesmo estatuto que o inglês, francês ou qualquer outro, sendo, assim, utilizada e reconhecida em seu país de origem (GALA, 2022).

3 ASSISTENTES VIRTUAIS

Os assistentes virtuais são uma das tecnologias mais fascinantes e revolucionárias da era digital. Eles são programas de computador que utilizam Inteligência Artificial para auxiliar os usuários na realização de tarefas, desde comandos simples, como agendar um compromisso, até tarefas mais complexas, como gerenciar tarefas e projetos. Com a popularidade crescente de dispositivos móveis e assistentes domésticos, a demanda por assistentes virtuais tem aumentado rapidamente. Esses programas estão se tornando cada vez mais sofisticados e integrados em nossas vidas cotidianas, transformando a maneira como interagimos com a tecnologia. Eles estão mudando a maneira como interagimos com a tecnologia, tornando a nossa vida mais fácil e eficiente. Neste capítulo, será explorado o mundo dos assistentes virtuais e o impacto que eles estão tendo na sociedade.

3.1 TECNOLOGIAS DE RECONHECIMENTO DE VOZ

O reconhecimento de voz é uma das tecnologias que mais têm evoluído nos últimos anos, graças ao grande interesse em aplicações de voz como assistentes virtuais, *chatbots*, sistemas de transcrição de fala para texto, sistemas de controle de dispositivos e muitas outras aplicações. As tecnologias de reconhecimento de voz se tornaram cada vez mais populares em várias indústrias, incluindo saúde, finanças, transporte e varejo, devido à sua capacidade de fornecer respostas mais rápidas e precisas, além de melhorar a experiência do usuário (BIBLIO FORA DA CAIXA, 2023).

A capacidade de reconhecer a voz é uma tecnologia baseada em Inteligência Artificial que permite a transcrição da fala humana em texto ou comando, sendo um processo complexo que envolve várias etapas, desde a gravação do áudio até o reconhecimento do conteúdo da fala. Para que o sistema funcione corretamente, é necessário que haja uma alta qualidade do áudio e que o modelo de reconhecimento de voz seja treinado com um grande volume de dados de fala em diversas línguas e dialetos.

As tecnologias que possuem essas características têm duas principais abordagens: a abordagem baseada em regras e a abordagem baseada em aprendizado de máquina. A primeira abordagem é mais antiga e consiste em

programar o sistema com regras específicas para reconhecer determinados padrões de fala. Já a abordagem baseada em aprendizado de máquina utiliza algoritmos de Inteligência Artificial que analisam e aprendem a partir de grandes quantidades de dados de fala (CYRINO, 2021).

Existem vários tipos de tecnologias de reconhecimento de voz, desde sistemas embarcados em dispositivos móveis, como *smartphones*, até sistemas de reconhecimento de fala remotos, que são acessados por meio de uma conexão com a internet. Entre os principais desenvolvimentos na área, estão o reconhecimento de voz *off-line*, que permite o processamento de fala sem conexão com a internet, e o contextual, que utiliza informações do contexto para melhorar a precisão do reconhecimento.

Apesar de todas as vantagens, ainda há desafios a serem superados nas tecnologias de reconhecimento de voz. Por exemplo, a qualidade do áudio é um fator crítico para o funcionamento correto do sistema, o que pode ser afetado por ruídos ambientais e sotaques regionais. Além disso, há preocupações com a privacidade dos usuários, especialmente quando os dispositivos com tecnologia de reconhecimento de voz estão sempre ouvindo e podem gravar conversas pessoais.

Em termos de perspectivas futuras, espera-se que as tecnologias de reconhecimento de voz continuem avançando e se tornando cada vez mais integradas ao nosso cotidiano, proporcionando uma melhor experiência do usuário e maior eficiência em diversas áreas.

3.1.1 Empresas

O reconhecimento de voz é uma tecnologia que tem ganhado cada vez mais espaço em grandes empresas, devido à sua capacidade de automatizar tarefas e melhorar a experiência do usuário. A Google, por exemplo, oferece o Google Assistant, um assistente virtual que pode ser controlado por voz em dispositivos móveis e em dispositivos inteligentes para casa. A Amazon, por sua vez, oferece a Alexa, que é capaz de controlar dispositivos inteligentes, fornecer informações e reproduzir música, entre outras funcionalidades (INBOT, 2023).

A Microsoft desenvolveu a Cortana, assistente virtual disponível em dispositivos com o sistema operacional Windows, enquanto a Apple oferece a Siri, assistente presente em dispositivos móveis e computadores com o sistema operacional macOS.

A Nuance é uma empresa que se destaca na área de reconhecimento de voz, oferecendo soluções para empresas em diversas áreas, incluindo saúde e finanças. A International Business Machines (IBM) também tem investido em tecnologias de reconhecimento de voz, oferecendo soluções de Inteligência Artificial para empresas em diversas áreas.

As tecnologias de reconhecimento de voz oferecem vantagens para as empresas que as utilizam, como a automatização de tarefas, a melhoria da eficiência e a redução de custos. Além disso, essas tecnologias melhoram a experiência do usuário, permitindo que as pessoas interajam com dispositivos e serviços de forma mais natural e intuitiva.

3.1.2 Saúde

Uma das principais aplicações dessa tecnologia na área da saúde é na criação de assistentes virtuais que auxiliam os médicos na coleta e análise de dados dos pacientes. Esses assistentes podem ser treinados para reconhecer a fala do médico e interpretar as informações coletadas, fornecendo sugestões de diagnóstico e tratamento. Isso torna o processo mais rápido e eficiente, permitindo que os médicos dediquem mais tempo ao atendimento de pacientes (LANGSTON, 2020).

Além disso, o reconhecimento de voz também pode ser aplicado em sistemas de transcrição de fala para texto em tempo real durante consultas médicas. Esses sistemas permitem que as informações coletadas durante as consultas sejam transcritas de forma precisa, evitando erros de interpretação e perda de informações importantes. Isso torna o processo mais confiável e ajuda a garantir a qualidade do atendimento.

Outra aplicação relevante é na criação de sistemas de monitoramento de pacientes em tempo real. Esses sistemas permitem que os médicos monitorem a condição de saúde dos pacientes à distância, com base nas informações coletadas pelo sistema de reconhecimento de voz. Isso torna o processo mais eficiente,

permitindo que os médicos identifiquem precocemente problemas de saúde e adotem medidas preventivas.

Como exemplos de ferramentas já existentes para esses objetivos, pode-se citar o Dragon Medical One, sistema desenvolvido pela Nuance que permite a transcrição de fala para texto em tempo real durante consultas médicas e a plataforma de assistente virtual para a saúde, que permite a criação de assistentes virtuais personalizados para atender às necessidades específicas de cada hospital ou clínica, criada pela Microsoft (LANGSTON, 2020).

3.1.3 Segurança

O reconhecimento de voz tem sido aplicado em diversas áreas, incluindo a área de sistemas de segurança, para controle de acesso e monitoramento. Essa tecnologia pode ser utilizada para autenticação de identidade de indivíduos em sistemas de segurança, como por exemplo, para desbloquear portas ou acessar informações restritas. Além disso, o reconhecimento de voz também pode ser utilizado em sistemas de monitoramento, permitindo a detecção de sons específicos e alertando para possíveis ameaças.

Um exemplo de aplicação de reconhecimento de voz na área de segurança é o sistema de controle de acesso da empresa francesa Safran Identity & Security, que utiliza o reconhecimento de voz para autenticar a identidade de indivíduos. O sistema permite o registro da voz de cada indivíduo autorizado e, em seguida, compara a voz do indivíduo que tenta acessar o local com as vozes registradas. O sistema é capaz de identificar a voz mesmo em ambientes ruidosos e tem uma taxa de erro extremamente baixa (SAFRAN, *sd*).

Outro exemplo é o uso de reconhecimento de voz em sistemas de monitoramento de segurança, como o sistema desenvolvido pela empresa americana Audio Analytics. O sistema utiliza Inteligência Artificial para analisar os sons captados por microfones e identificar padrões específicos de sons que podem indicar atividades suspeitas, como a quebra de vidro, barulhos de passos ou até mesmo tiros. O sistema é capaz de alertar imediatamente as autoridades para possíveis ameaças (TECH.EU, 2022).

O uso de reconhecimento de voz na área de sistemas de segurança tem se mostrado uma solução eficiente para aumentar a segurança de ambientes e garantir a autenticação de identidade de indivíduos. No entanto, é importante ressaltar que a privacidade das informações pessoais deve ser levada em consideração ao implementar essas tecnologias.

3.1.4 Atendimento em *call centers*

A qualidade do atendimento em *call centers* é um tema crucial para empresas que buscam fornecer um serviço de qualidade aos seus clientes. Eles são os principais canais de contato entre clientes e empresas, e por isso, é fundamental que o atendimento seja eficiente e eficaz.

Diversas estratégias são utilizadas para melhorar a qualidade do atendimento em *call centers*, incluindo a implementação de treinamentos específicos para os agentes, a adoção de tecnologias que permitam um atendimento mais rápido e personalizado, a realização de pesquisas de satisfação do cliente e a criação de processos internos eficientes. Com o uso da tecnologia de reconhecimento de voz, é possível realizar a identificação automática do cliente e a transcrição da fala, o que permite uma resposta mais rápida e precisa (VOCALCOM, *sd*).

A implementação de soluções de reconhecimento de voz em *call centers* pode ajudar a melhorar a experiência do cliente, reduzir o tempo de espera para atendimento e aumentar a eficiência do serviço prestado. No entanto, é importante lembrar que a qualidade do atendimento não depende apenas de tecnologias avançadas, mas também de uma equipe bem treinada e processos internos bem definidos.

3.2 TRADUÇÃO DE TEXTO PARA LINGUAGEM DE SINAIS

As tecnologias de tradução de texto para linguagem de sinais têm sido desenvolvidas para ajudar pessoas surdas ou com deficiência auditiva a acessar informações e se comunicar mais facilmente em situações cotidianas. Essas tecnologias utilizam algoritmos de Processamento de Linguagem Natural (PLN) para traduzir o texto em uma representação visual de sinais (SANCHEZ, *sd*).

Existem diferentes tipos de tecnologias de tradução de texto para linguagem de sinais, como *softwares* e aplicativos de reconhecimento de voz, que transformam a fala em texto para, em seguida, traduzi-lo em linguagem de sinais, e *softwares* que fazem a tradução diretamente do texto para a linguagem de sinais. Além disso, existem também tecnologias que permitem a interação em tempo real, por meio de videoconferência, entre pessoas surdas e ouvintes, com o auxílio de um intérprete de linguagem de sinais.

A tradução de texto para linguagem de sinais pode ser realizada por meio de vídeos, animações ou avatares animados que reproduzem os movimentos das mãos, expressões faciais e corporais dos intérpretes de língua de sinais. Essas tecnologias têm sido aplicadas em diversas áreas, como educação, saúde, negócios e entretenimento.

Um exemplo de tecnologia para linguagem de sinais é o projeto Hand Talk, desenvolvido no Brasil. O aplicativo utiliza Inteligência Artificial para traduzir textos em português para a Língua Brasileira de Sinais, por meio de um avatar animado que representa um intérprete de Libras, como visto na Figura 1. O Hand Talk já foi premiado internacionalmente e é utilizado em diversas instituições, como escolas e hospitais, para promover a acessibilidade e a inclusão social de pessoas com deficiência auditiva (VOCÊ S/A, 2021).

Figura 1 - Hand Talk



Fonte: HANDTALK, *sd, online.*

No entanto, ainda existem desafios para a efetividade das tecnologias de tradução de texto para linguagem de sinais, como a necessidade de treinamento dos modelos de reconhecimento de fala e gestos, a falta de padronização dos sinais em diferentes regiões e línguas, e a disponibilidade de recursos financeiros e tecnológicos para a implantação dessas tecnologias em larga escala.

Apesar dos desafios, as tecnologias de tradução de texto para linguagem de sinais têm apresentado avanços significativos, contribuindo para a melhoria da comunicação entre pessoas surdas e ouvintes em diferentes contextos, como na educação, no mercado de trabalho e na saúde.

Algumas empresas e organizações têm investido em tecnologias de tradução de texto para linguagem de sinais, como a Microsoft, que desenvolveu um aplicativo de tradução de texto para linguagem de sinais, e a Federação Mundial de Surdos, que lançou um dicionário *online* de sinais em diversas línguas (MICROSOFT, 2022).

Além disso, governos têm implementado políticas de acessibilidade e inclusão, que incluem a utilização de tecnologias de tradução de texto para linguagem de sinais. Dentre as principais medidas com o intuito de melhorar a qualidade de vida dessa parcela da população, é importante ressaltar a primeira lei totalmente voltada à acessibilidade e a mais importante, a Lei Nº 10.098, que segundo a ementa oficial:

Art. 1º Esta Lei estabelece normas gerais e critérios básicos para a promoção da acessibilidade das pessoas portadoras de deficiência ou com mobilidade reduzida, mediante a supressão de barreiras e de obstáculos nas vias e espaços públicos, no mobiliário urbano, na construção e reforma de edifícios e nos meios de transporte e de comunicação (BRASIL, 2000, p. 2).

Vale mencionar que fica a critério de órgão públicos como o Ministério Público Federal, aliado as Secretarias da Pessoa com Deficiência a fiscalização e cumprimento das leis e diretrizes que regem os direitos das pessoas com deficiência.

3.2.1 Processamento de Linguagem Natural

O Processamento de Linguagem Natural é um campo da ciência da computação que busca desenvolver algoritmos capazes de compreender e produzir linguagem humana, de forma que computadores possam interagir com seres humanos em sua própria linguagem. Na tradução de texto para linguagem de sinais,

ele é utilizado para desenvolver tecnologias capazes de interpretar e traduzir o texto em linguagem natural para uma representação visual de sinais.

As tecnologias de tradução de texto para linguagem de sinais utilizam algoritmos de PLN para reconhecer e interpretar a linguagem natural contida em textos escritos ou falados. Esses algoritmos são treinados com milhares de exemplos de textos e sinais, de forma que possam aprender a reconhecer padrões e identificar a correspondência entre eles. O Processamento de Linguagem Natural permite ainda que as tecnologias de tradução de texto para linguagem de sinais sejam adaptáveis a diferentes línguas e regiões, uma vez que podem ser treinadas com diferentes conjuntos de dados (SANCHEZ, *sd*).

No entanto, a aplicação dele na tradução de texto para linguagem de sinais ainda apresenta desafios, como a necessidade de reconhecimento de gestos, expressões faciais e corporais, que exigem a utilização de tecnologias de reconhecimento de imagem e vídeo. Além disso, a compreensão de nuances da linguagem humana, como ironia e sarcasmo, ainda é um desafio para os algoritmos do processamento.

Independente das adversidades, ele tem se mostrado uma ferramenta fundamental para o desenvolvimento de tecnologias de tradução de texto para linguagem de sinais mais efetivas e acessíveis. Empresas como a Microsoft e a Google têm investido em pesquisa e desenvolvimento de tecnologias de PLN aplicadas à tradução de texto para linguagem de sinais, contribuindo para a melhoria da comunicação e inclusão social de pessoas com deficiência auditiva.

3.2.2 Hand Talk

Aplicativo brasileiro de tradução de texto para linguagem de sinais, o Hand Talk utiliza Inteligência Artificial para traduzir textos em português para a Língua Brasileira de Sinais, por meio de um avatar animado que representa um intérprete de LIBRAS. O projeto foi desenvolvido em 2012, em Recife, pelos empreendedores Ronaldo Tenório, Thadeu Luz e Carlos Wanderlan (VOCÊ S/A, 2021).

Ele utiliza tecnologia de reconhecimento de voz e gestos para traduzir textos em português para a LIBRAS, por meio de um avatar animado que representa um intérprete de língua de sinais. O avatar realiza os sinais em tempo real e pode ser

personalizado pelo usuário de acordo com suas preferências. Além disso, oferece uma plataforma para a criação de conteúdo em LIBRAS, que pode ser compartilhado em redes sociais ou incorporado em sites.

O Hand Talk já recebeu diversos prêmios internacionais, como o Prêmio de Inovação da ONU, em 2013, e o Prêmio de Empreendedorismo Social da Fundação Schwab, em 2015. Além disso, o aplicativo já foi utilizado em eventos de grande porte, como a cerimônia de abertura dos Jogos Olímpicos Rio 2016.

No entanto, ele ainda enfrenta desafios, como a falta de padronização dos sinais em diferentes regiões e línguas, e a disponibilidade de recursos financeiros e tecnológicos para a implantação do aplicativo em larga escala.

3.2.3 FLAC

O *Free Lossless Audio Codec* (FLAC) é um formato de arquivo usado para armazenar dados de áudio digital. Como outros formatos de arquivo de áudio, como *MPEG-1/2 Audio Layer 3* (MP3) ou *Waveform Audio File Format* (WAV), os arquivos FLAC são de código aberto e compactam os dados de áudio do usuário. No entanto, ele é um tipo de arquivo sem perdas, o que lhe confere uma vantagem significativa sobre outros tipos de arquivos de áudio (SANTOS, *sd*).

Os arquivos FLAC preservam todos os dados e informações do áudio durante o processo de codificação e compactação, o que resulta em uma resolução de áudio impressionantemente alta. Em comparação com outros formatos, fornecem melhores taxas de amostragem e faixas mais amplas de profundidade. Por esse motivo, oferecem melhor qualidade de som e são preferidos por usuários em todo o mundo. No entanto, esses detalhes codificados no arquivo FLAC acabam aumentando o tamanho do arquivo em comparação com outros formatos de arquivo de áudio.

Outra característica importante do formato de arquivo FLAC é o suporte multiplataforma. Anteriormente, eles não eram suportados em muitas plataformas, mas com atualizações recentes, quase todas as plataformas e reprodutores multimídia estenderam o suporte para esse formato de arquivo de áudio. Esse formato é suportado pelo Android 3.1 em diante e pelo iPhone 5S e superior, além dos sistemas operacionais Windows, Mac e Linux. Até mesmo a maioria dos alto-falantes portáteis e estéreos do carro agora pode reproduzi-los sem problemas.

Uma vantagem considerável que possuem é a sua alta velocidade de decodificação. Mesmo em um sistema modesto, pode-se obter desempenho de decodificação em tempo real com facilidade e rapidez, e são resistentes a erros de *streaming* e são menos propensos a serem corrompidos. Mesmo que ocorra um erro, somente uma pequena parte do áudio será afetada, geralmente limitada ao quadro ou a uma fração de segundo em que ocorreu. Em outros tipos de arquivos de áudio sem perdas, um único erro normalmente destrói todo o fluxo.

O fato de ser um formato de arquivo de áudio de código aberto é o que o torna perfeito para necessidades de arquivamento, oferecendo assim uma opção de verificação que funciona para decodificar o fluxo simultaneamente enquanto o processo de codificação está ocorrendo, permitindo que realize a comparação de duas versões e até mesmo aborte o processo em caso de incompatibilidade.

Ao compará-lo com outros formatos populares de arquivo de áudio, como MP3 ou WAV, há diferenças significativas, se tratando do MP3 é um formato com perdas, o que significa que a qualidade do áudio é comprometida durante a compressão, enquanto o formato WAV é um tipo de arquivo sem perdas, mas não é tão eficiente em termos de tamanho de arquivo quanto o FLAC (SANTOS, *sd*).

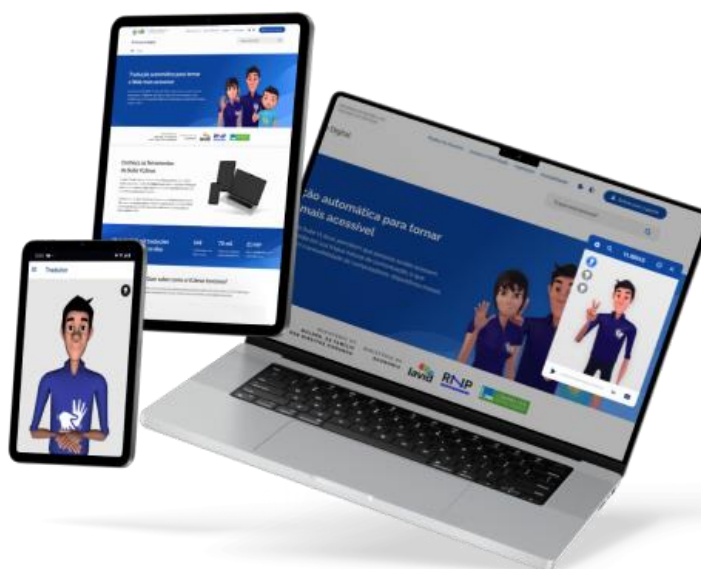
Por fim, o projeto CIDA, que tem como objetivo tornar a informação mais acessível para pessoas com deficiência auditiva, utiliza o FLAC para tradução de textos para linguagem de sinais, pois é necessário que a imagem e som estejam sincronizados para que a informação seja compreendida de forma clara e precisa. Dessa forma, o FLAC é uma opção ideal para a tradução, pois permite que o áudio seja compactado para um tamanho menor, facilitando a transmissão do arquivo pela internet e garantindo que a qualidade do som seja mantida, além do fato que o formato livre de patentes permite ser usado por qualquer pessoa sem a necessidade de pagar *royalties*.

3.3 DESENVOLVIMENTO DE ASSISTENTES VIRTUAIS

Os assistentes virtuais são programas de computador que utilizam Inteligência Artificial para interagir com usuários de forma natural e personalizada, ajudando-os em tarefas diversas, como mostra a Figura 2. Esses sistemas são capazes de compreender a linguagem natural, interpretar solicitações e oferecer respostas em

tempo real. Os assistentes virtuais podem ser encontrados em diversos dispositivos, como *smartphones*, *tablets*, *smart speakers* e outros dispositivos *Internet of Things* (IoT) (INBOT, 2023).

Figura 2 – Assistentes Virtuais



Fonte: ASSISTENTES, 2023, *online*.

Os assistentes virtuais são projetados para aprender com as interações dos usuários, o que significa que sua capacidade de entender e responder a perguntas e comandos melhora com o tempo.

O desenvolvimento de assistentes virtuais requer uma série de etapas, desde a coleta e organização dos dados até a criação de modelos de linguagem e processamento de fala. É necessário que o sistema seja treinado com um grande volume de dados para que possa oferecer respostas precisas e personalizadas. Além disso, é importante que o assistente virtual seja capaz de se adaptar a diferentes contextos e interações (INBOT, 2023).

As tecnologias utilizadas no desenvolvimento de assistentes virtuais incluem algoritmos de aprendizado de máquina, processamento de linguagem natural, processamento de fala e reconhecimento de voz. Essas tecnologias estão em constante evolução, permitindo que os assistentes virtuais se tornem cada vez mais inteligentes e eficientes.

O desenvolvimento de assistentes virtuais tem um grande potencial para transformar a maneira como interagimos com a tecnologia e com as empresas. Com

a crescente adoção de dispositivos inteligentes e a proliferação de dados, os assistentes virtuais têm a capacidade de fornecer uma experiência de usuário mais personalizada e eficiente.

3.3.1 Modelos de linguagem

Modelos de linguagem são componentes essenciais no desenvolvimento de assistentes virtuais capazes de interpretar a linguagem natural dos usuários. Esses modelos são algoritmos de aprendizado de máquina treinados com grandes quantidades de dados de texto, que aprendem a prever a próxima palavra ou frase com base no contexto (DEEP LEARNING BOOK, *sd*).

Eles permitem que os assistentes virtuais entendam a intenção do usuário, mesmo em casos de solicitações ambíguas ou incompletas, melhorando a precisão das respostas. Além disso, eles são capazes de aprender com as interações dos usuários, aprimorando ainda mais a capacidade de interpretar e responder às solicitações.

Existem diferentes tipos de modelos de linguagem, como os baseados em redes neurais, que utilizam técnicas como o *Transformer* e o Bidirectional Encoder Representations from *Transformers* (BERT) para melhorar a compreensão do contexto e aprimorar a capacidade de resposta dos assistentes virtuais.

3.3.2 Redes Neurais

As redes neurais são um tipo de algoritmo que imita o funcionamento do cérebro humano para resolver problemas complexos. Esses algoritmos são capazes de aprender com os dados e encontrar padrões em conjuntos de informações.

Segundo Géron (2019), as redes neurais são compostas por camadas de neurônios que processam informações e as transmitem para as camadas seguintes. Cada neurônio realiza uma operação matemática simples e a saída desse neurônio é enviada para outros neurônios da camada seguinte. O processo de treinamento de uma rede neural consiste em ajustar os pesos das conexões entre os neurônios para que o algoritmo seja capaz de identificar padrões nos dados.

As redes neurais são amplamente utilizadas em áreas como reconhecimento de imagens, processamento de linguagem natural, previsão de séries temporais e classificação de dados. Uma das vantagens desses algoritmos é sua capacidade de lidar com grandes quantidades de dados e encontrar padrões em meio a ruídos e variações (IBM, *sd*).

Além disso, as redes neurais são capazes de se adaptar a novos dados e situações, o que as torna extremamente úteis em ambientes em constante mudança. De acordo com Haykin (2001), as redes neurais são uma das áreas mais promissoras da Inteligência Artificial e têm potencial para transformar completamente a forma como lidamos com dados e informações.

3.3.3 Aprendizado de Máquina

O aprendizado de máquina é uma das principais tecnologias utilizadas no desenvolvimento de assistentes virtuais. Essa tecnologia permite que os sistemas aprendam com dados e interações anteriores, adaptando-se às necessidades e preferências dos usuários ao longo do tempo.

Existem diversos algoritmos de aprendizado de máquina que podem ser aplicados no desenvolvimento de assistentes virtuais, como redes neurais, árvores de decisão, algoritmos genéticos, entre outros. Cada algoritmo tem suas vantagens e desvantagens, e a escolha depende do tipo de problema que está sendo resolvido e dos dados disponíveis para treinamento (CYRINO, 2021).

No caso dos assistentes virtuais, o aprendizado é utilizado para criar modelos de linguagem e processamento de fala. Isso permite que o sistema entenda a linguagem natural utilizada pelos usuários e forneça respostas precisas e personalizadas. Além disso, o aprendizado de máquina é utilizado para melhorar a capacidade do assistente virtual de reconhecer padrões e identificar contextos diferentes.

3.3.4 Processamento de fala em assistentes virtuais

O processamento de fala é uma tecnologia utilizada em assistentes virtuais para que eles possam reconhecer e interpretar a fala humana. Com a evolução dos

sistemas de reconhecimento de voz, os assistentes virtuais estão cada vez mais precisos em compreender o que o usuário está falando e oferecer respostas adequadas (BIBLIO FORA DA CAIXA, 2023).

Para que um assistente virtual seja capaz de processar a fala, é necessário que ele tenha acesso a um grande banco de dados de fala humana, para que possa entender as variações de sotaque, entonação e ritmo da fala. Além disso, o sistema deve ser capaz de filtrar ruídos externos e identificar a voz do usuário.

O processamento é uma tecnologia complexa, que envolve diversas etapas, como o reconhecimento da fala, a separação dos sons em fonemas, a identificação da entonação e a tradução em texto. É importante que essas etapas sejam realizadas de forma rápida e precisa para que o assistente virtual possa oferecer uma resposta em tempo real (SANCHEZ, *sd*).

Alguns dos desafios enfrentados em assistentes virtuais incluem a identificação de palavras homófonas (palavras que possuem o mesmo som, mas significados diferentes), o reconhecimento de fala em ambientes ruidosos e a adaptação a diferentes sotaques e dialetos.

Apesar dos desafios, o processamento de fala é uma tecnologia em constante evolução, e cada vez mais presente em nossas vidas por meio de assistentes virtuais, citados anteriormente, como a Siri, Alexa e Google Assistant.

4 LEVANTAMENTO DE REQUISITOS E FERRAMENTAS

O processo de desenvolvimento de um *software* é uma tarefa complexa e desafiadora. Uma das etapas fundamentais nesse processo é o levantamento de requisitos, que envolve a identificação e documentação das necessidades e expectativas do cliente em relação ao *software* a ser desenvolvido. O sucesso do *software* depende em grande parte da precisão e qualidade do levantamento de requisitos, pois isso determina a viabilidade e a eficácia do produto.

Neste capítulo, será explorado o processo de levantamento de requisitos em detalhes, discutindo as ferramentas e técnicas utilizadas para coletar, analisar e documentar os requisitos. Além disso, serão apresentadas as melhores práticas para garantir a qualidade dos requisitos e o sucesso do projeto.

4.1 MATRIZ SWOT

A análise de *Strengths, Weaknesses, Opportunities e Threats* (SWOT) é uma ferramenta utilizada no planejamento estratégico de projetos e empresas. A sigla SWOT representa os termos em inglês: forças, fraquezas, oportunidades e ameaças respectivamente. Como é possível observar na Figura 3, por meio da análise desses quatro elementos, é possível identificar pontos fortes e fracos do projeto, bem como oportunidades e ameaças que podem afetar o seu desenvolvimento.

Figura 3 - Matriz SWOT

<p>Strengths</p> <ul style="list-style-type: none"> • Solução tecnológica inovadora e única no mercado, com grande potencial de impacto social. • Capaz de converter o áudio falado em tempo real em texto desejado, o que torna o processo mais rápido e eficiente. • Possibilidade de integração com assistentes virtuais para reprodução da linguagem de sinais, ampliando a acessibilidade dos deficientes auditivos. 	<p>Weaknesses</p> <ul style="list-style-type: none"> • Necessidade de infraestrutura tecnológica robusta para funcionamento da solução, o que pode gerar custos elevados para implementação. • Dependência de conexão com a internet para funcionamento adequado da solução, o que pode limitar a utilização em locais com baixa conectividade. • Requer treinamento para utilização, tanto por parte dos deficientes auditivos quanto dos profissionais responsáveis pela implementação e manutenção da solução.
<p>Opportunities</p> <ul style="list-style-type: none"> • Mercado em crescimento na área de tecnologia assistiva para pessoas com deficiência, o que pode atrair investimentos e parcerias estratégicas. • Potencial de expansão do projeto para outros idiomas e países, ampliando o impacto social e possibilitando a internacionalização da solução. • Possibilidade de integração com outras soluções tecnológicas voltadas para acessibilidade, criando um ecossistema de soluções interconectadas. 	<p>Threats</p> <ul style="list-style-type: none"> • Possibilidade de surgimento de concorrentes com soluções semelhantes, o que pode gerar competitividade no mercado. • Instabilidade política e econômica do país pode afetar o investimento em tecnologia e projetos sociais. • Barreiras culturais e sociais que ainda existem em relação às pessoas com deficiência podem limitar a adoção e utilização da solução.

Fonte: Autoria própria.

4.2 5W2H

A ferramenta 5W2H é uma metodologia simples e eficaz que tem como objetivo auxiliar na elaboração e execução de um plano de ação. O significado do acrônimo 5W2H pode ser visualizado no Quadro 2.

Quadro 2 – Componentes 5W2H

Componentes	Descrição
<i>What</i>	O que deve ser feito?
<i>Why</i>	Por que deve ser feito?
<i>Who</i>	Quem será responsável por fazer?
<i>Where</i>	Onde será realizado?
<i>When</i>	Quando será realizado?
<i>How</i>	Como será feito?
<i>How much</i>	Quanto custará?

Fonte: Autoria própria

Essas perguntas ajudam a definir com clareza as ações a serem tomadas, quem serão os responsáveis e como elas serão executadas. A metodologia 5W2H é especialmente útil para projetos que envolvem múltiplas etapas ou equipes, e ajuda a garantir que todos os envolvidos estejam alinhados quanto aos objetivos e prazos. O Quadro 3 apresenta sua utilização em relação ao desenvolvimento da ferramenta CIDA.

Quadro 3 - 5W2H

PERGUNTA	RESPOSTA
WHAT	Desenvolvimento da ferramenta CIDA (conversor instantâneo para deficientes auditivos).
WHY	Oferecer um recurso de acessibilidade eficaz e abrangente para a população de deficientes auditivos, permitindo-lhes consumir conteúdos audiovisuais de forma mais e eficiente.
WHERE	Navegadores
WHEN	Desenvolvimento do software: 6 meses. Testes: 1 mês. Finalização e lançamento do produto: 1 mês.
WHO	Equipe de desenvolvimento de software e profissionais em acessibilidade
HOW	Desenvolvimento do software utilizando linguagens de programação como Python. Testes realizados em ambiente controlado com a participação de pessoas com deficiência auditiva. Lançamento do produto por meio de campanhas de marketing digital e parcerias com organizações que atuam com deficientes auditivos.
HOW MUCH	Sem custo

Fonte: Autoria própria

4.3 BPMN

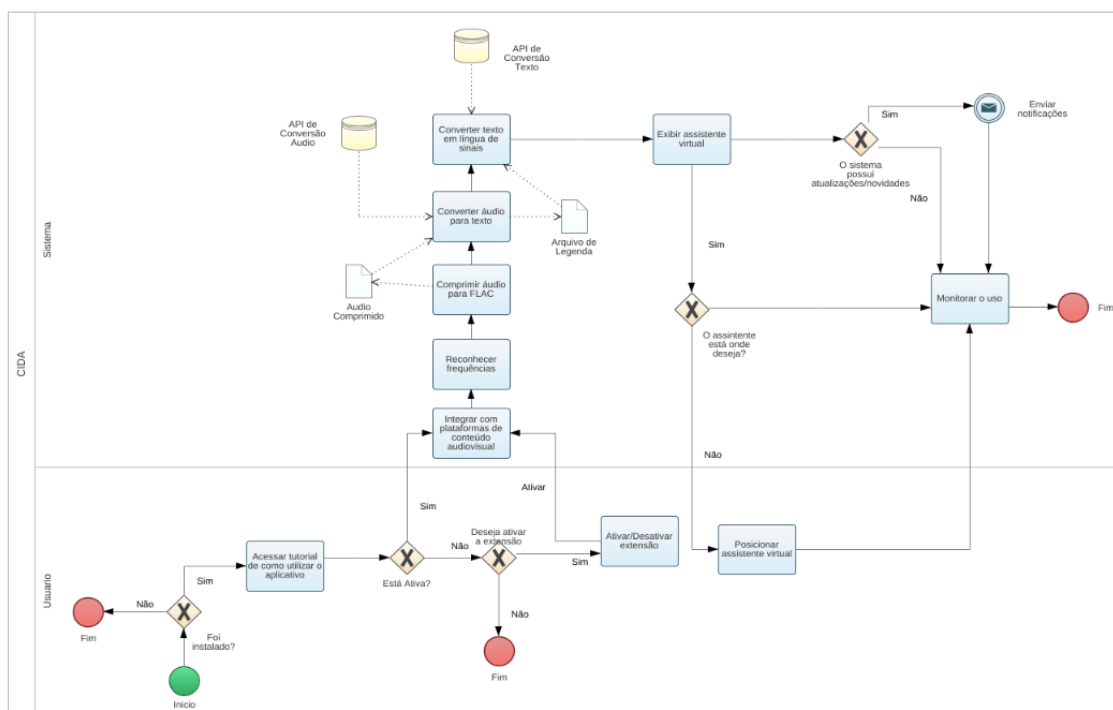
Business Process Model and Notation (BPMN) é uma notação gráfica padronizada para modelar processos de negócios. É uma linguagem visual que permite representar de maneira clara e compreensível como um processo é executado, desde o início até o fim.

O BPMN é composto por um conjunto de símbolos e notações que representam as atividades, eventos, fluxos, decisões, *gateways*, objetos de dados e participantes

envolvidos em um processo de negócio. Ele permite que os usuários criem diagramas que ilustram claramente como as informações e as atividades se relacionam em um processo, mostrando também as decisões, ramificações e loops que ocorrem ao longo do caminho.

Com o BPMN, é possível modelar processos de negócios com diferentes níveis de detalhamento, desde a visão macro do processo até o nível mais granular. Além disso, o BPMN permite a comunicação e o alinhamento entre diferentes partes interessadas em um processo de negócio, incluindo gestores, analistas de negócios, desenvolvedores e usuários finais. A padronização do BPMN também facilita a integração entre diferentes sistemas de informação e ajuda a garantir a conformidade com as melhores práticas de gerenciamento de processos de negócios. Na Figura 4 é utilizada essa modelagem para identificar o fluxo do sistema proposto.

Figura 4 – BPMN



Fonte: Autoria própria

4.4 REQUISITOS FUNCIONAIS

A especificação de requisitos funcionais é uma etapa fundamental no desenvolvimento de um *software*, pois esses requisitos são a descrição detalhada das funções que o *software* deve executar e dos resultados que ele deve produzir. Isso inclui todas as interações que o *software* deve ter com seus usuários, bem como as respostas que o *software* deve fornecer em relação às entradas do usuário.

No Quadro 4 está apresentada a especificação de requisitos da CIDA, que é a base para o *design*, desenvolvimento, testes e manutenção do *software*, e deve ser elaborada com muito cuidado para garantir que todas as necessidades do usuário sejam atendidas.

Quadro 4 - Requisitos Funcionais

ID:	RF001 - Acessar tutorial de como utilizar o aplicativo
Descrição:	Ao instalar a extensão, deve ser aberta uma aba para redirecionamento a um vídeo tutorial mostrando como utilizar a ferramenta.
Categoria:	() Oculto (X) Evidente
Prioridades:	() Essencial () Alta (X) Média () Baixa
Regra de Negócio:	RN005 – Compatibilidade com navegadores
Atores:	Usuário
ID:	RF002 - Integrar com plataformas de conteúdo audiovisual
Descrição:	O sistema deve ser capaz de se integrar com diversas plataformas de conteúdo audiovisual (como Youtube, Vimeo, Netflix etc.) para capturar o áudio falado e converter em texto em tempo real.
Categoria:	(X) Oculto () Evidente
Prioridades:	(X) Essencial () Alta () Média () Baixa
Regra de Negócio:	RN004 – Compatibilidade com plataformas de <i>streaming</i>
Atores:	Sistema
ID:	RF003 – Reconhecer frequências
Descrição:	A ferramenta deve reconhecer frequências que representem uma fala no áudio que está sendo reproduzido para eventualmente comprimi-lo.

Categoria:	(X) Oculto () Evidente
Prioridades:	(X) Essencial () Alta () Média () Baixa
Regra de Negócio:	RN002 – <i>Open Source</i>
Atores:	Sistema
ID:	RF004 – Comprimir áudio para FLAC
Descrição:	O sistema deve ser capaz de capturar o áudio falado em <i>streams</i> e converter em texto para posterior conversão em linguagem de sinais.
Categoria:	(X) Oculto () Evidente
Prioridades:	(X) Essencial () Alta () Média () Baixa
Regra de Negócio:	RN002 – <i>Open Source</i>
Atores:	Sistema
ID:	RF005 – Converter áudio para texto
Descrição:	O sistema deve ser capaz de enviar as frequências compactadas para uma API que irá converter em texto para posterior conversão em linguagem de sinais.
Categoria:	(X) Oculto () Evidente
Prioridades:	(X) Essencial () Alta () Média () Baixa
Regra de Negócio:	RN002 – <i>Open Source</i>
Atores:	Sistema
ID:	RF006 - Posicionar assistente virtual
Descrição:	O assistente virtual poderá ser posicionado em qualquer canto da tela, alterando sua exibição clicando em um ícone localizado perto do assistente.
Categoria:	() Oculto (X) Evidente
Prioridades:	() Essencial () Alta (X) Média () Baixa
Regra de Negócio:	RN004 – Compatibilidade com plataformas de <i>streaming</i> RN005 – Compatibilidade com navegadores
Atores:	Usuário

ID:	RF007 – Converter texto em língua de sinais
Descrição:	A ferramenta deve enviar os dados resultantes da conversão de áudio para texto para uma biblioteca que retornará o assistente virtual realizando os gestos
Categoria:	(X) Oculto () Evidente
Prioridades:	(X) Essencial () Alta () Média () Baixa
Regra de Negócio:	RN002 – <i>Open Source</i>
Atores:	Sistema
ID:	RF008 – Exibir assistente virtual
Descrição:	Para concluir o processo principal da ferramenta, após enviar o texto para API responsável por converter a legenda em gestos, será necessário exibir o assistente virtual que reproduzirá os movimentos.
Categoria:	() Oculto (X) Evidente
Prioridades:	(X) Essencial () Alta () Média () Baixa
Regra de Negócio:	RN002 – <i>Open Source</i> RN004 – Compatibilidade com plataformas de <i>streaming</i> RN005 – Compatibilidade com navegadores
Atores:	Sistema
ID:	RF009 – Monitorar o uso
Descrição:	O sistema deve ser capaz de monitorar o uso da ferramenta e gerar relatórios para análise da equipe de desenvolvimento e melhoria contínua da ferramenta.
Categoria:	(X) Oculto () Evidente
Prioridades:	() Essencial (X) Alta () Média () Baixa
Regra de Negócio:	RN001 – Manutenção RN004 – Compatibilidade com plataformas de <i>streaming</i> RN005 – Compatibilidade com navegadores
Atores:	Sistema
ID:	RF010 – Enviar notificações
Descrição:	A extensão deve enviar notificações para o usuário informando sobre atualizações e melhorias da ferramenta.
Categoria:	() Oculto (X) Evidente

Prioridades:	() Essencial (X) Alta () Média () Baixa
Regra de Negócio:	RN001 – Manutenção RN005 – Compatibilidade com navegadores
Atores:	Sistema
ID:	RF011 – Ativar/Desativar extensão
Descrição:	A extensão deve permitir que o usuário ative e desative a opção.
Categoria:	() Oculto (X) Evidente
Prioridades:	() Essencial (X) Alta () Média () Baixa
Regra de Negócio:	N/A
Atores:	Usuário

Fonte: Autoria própria.

4.5 REQUISITOS NÃO FUNCIONAIS

Ao contrário dos requisitos funcionais, que descrevem o comportamento do *software*, os requisitos não funcionais descrevem as propriedades e características do *software* que não estão diretamente relacionadas às suas funcionalidades. Esses requisitos podem incluir aspectos como desempenho, segurança, confiabilidade, facilidade de manutenção e outros atributos que são importantes para o sucesso do *software*. No Quadro 5, os aspectos tecnológicos fundamentais da plataforma proposta são elucidados.

A especificação cuidadosa dos requisitos não funcionais é essencial para garantir que o *software* atenda às necessidades e expectativas do usuário e seja capaz de operar efetivamente em seu ambiente de implantação.

Quadro 5 - Requisitos Não Funcionais

ID:	RNF001 - Linguagem de desenvolvimento
Restrição:	O sistema deverá ser construído utilizando a linguagem Python visando o melhor desempenho para o usuário e praticidade para o desenvolvimento.
Categoria:	(X) Oculto () Evidente

Obrigatoriedade:	() Desejável (X) Obrigatório
Permanência:	(X) Permanente () Transitório
ID:	RNF002 - Responsividade
Restrição:	O assistente virtual deverá se adaptar a extensão da tela em que está sendo exibido.
Categoria:	() Oculto (X) Evidente
Obrigatoriedade:	() Desejável (X) Obrigatório
Permanência:	(X) Permanente () Transitório
ID:	RNF003 - Idiomas de tradução
Restrição:	O sistema deverá suportar a linguagem de sinais conhecida como Língua Brasileira de Sinais (LIBRAS) como principal instrumento de comunicação.
Categoria:	() Oculto (X) Evidente
Obrigatoriedade:	(X) Desejável () Obrigatório
Permanência:	() Permanente (X) Transitório
ID:	RNF004 - Disponibilidade
Restrição:	O sistema deve estar disponível 24 horas por dia, 7 dias por semana, garantindo que o usuário possa acessar o conteúdo audiovisual convertido em tempo real sempre que precisar.
Categoria:	(X) Oculto () Evidente
Obrigatoriedade:	() Desejável (X) Obrigatório
Permanência:	(X) Permanente () Transitório
ID:	RNF005 - Interface de usuário intuitiva
Restrição:	Interface do usuário deve ser intuitiva e de fácil utilização, garantindo que qualquer pessoa possa utilizar a ferramenta sem dificuldades.
Categoria:	() Oculto (X) Evidente
Obrigatoriedade:	() Desejável (X) Obrigatório
Permanência:	(X) Permanente () Transitório

ID:	RNF006 – Uso em tempo real
Restrição:	A ferramenta deve ter um tempo de resposta rápido o suficiente para garantir que a tradução de áudio para texto e a conversão para LIBRAS ocorram em tempo real, sem atrasos perceptíveis para o usuário.
Categoria:	() Oculto (X) Evidente
Obrigatoriedade:	() Desejável (X) Obrigatório
Permanência:	(X) Permanente () Transitório
ID:	RNF007 – Acurácia na conversão de áudio para texto
Restrição:	É necessário garantir que a conversão de áudio para texto seja precisa e acurada, para que o resultado reproduzido na linguagem de sinais seja de qualidade e fácil compreensão para os usuários
Categoria:	(X) Oculto () Evidente
Obrigatoriedade:	(X) Desejável () Obrigatório
Permanência:	(X) Permanente () Transitório

Fonte: Autoria própria.

4.6 REGRAS DE NEGÓCIOS

As regras de negócios descrevem o conjunto de políticas, procedimentos e práticas que uma organização segue em suas operações diárias. Elas podem ser aplicadas em diversas áreas, como finanças, recursos humanos, *marketing*, entre outras.

O Quadro 6 mostra a especificação de regras de negócios. Essas regras são importantes para o desenvolvimento de um *software* que atenda às necessidades da organização. Ela envolve a identificação, documentação e comunicação dessas regras, garantindo que o *software* a ser desenvolvido seja capaz de suportá-las e implementá-las de forma eficaz.

Quadro 6 – Regras de Negócios

ID:	RN001 – Manutenção
Descrição:	O sistema deve estar sujeito a manutenção regular e atualizações de <i>software</i> para garantir a disponibilidade e eficácia contínuas da ferramenta.

ID:	RN002 – <i>Open Source</i>
Descrição:	Visando tornar essa ferramenta o mais acessível possível, é fundamental que sejam utilizadas majoritariamente bibliotecas e APIs <i>Open Sources</i> para o desenvolvimento (VLibras, ferramenta disponibilizada pelo Governo Brasileiro é um exemplo).
ID:	RN003 – Suporte técnico e atendimento ao usuário
Descrição:	O sistema deve contar com suporte técnico eficiente e atendimento ao usuário, para que possíveis problemas e dúvidas possam ser solucionados com rapidez e eficiência.
ID:	RN004 – Compatibilidade com plataformas de <i>streaming</i>
Descrição:	A extensão deve ser compatível com as principais plataformas de <i>streaming</i> de conteúdo audiovisual, como YouTube, Netflix, Amazon Prime, entre outras.
ID:	RN005 – Compatibilidade com navegadores
Descrição:	O sistema deve ser compatível com diferentes navegadores (como Google Chrome, Mozilla Firefox, Microsoft <i>Edge</i> etc.), garantindo uma experiência consistente para todos os usuários.

Fonte: Autoria própria.

4.7 MATRIZ DE RASTREABILIDADE (REQUISITOS FUNCIONAIS X REGRAS DE NEGÓCIO)

A matriz de rastreabilidade é uma ferramenta utilizada em gerenciamento de projetos para acompanhar e rastrear a implementação de requisitos em diferentes etapas do projeto. Ela consiste em uma tabela que relaciona requisitos específicos do projeto com as atividades e artefatos que foram criados para atender a esses requisitos, assim é possível garantir que todos os requisitos do projeto sejam atendidos e que todas as atividades necessárias para a sua implementação estejam sendo realizadas.

No Quadro 7, a matriz de rastreabilidade também ajuda a identificar possíveis lacunas ou inconsistências nos requisitos do projeto e a facilitar a comunicação entre as diferentes equipes envolvidas no projeto.

Quadro 7 - Matriz de Rastreabilidade – RF X RN

	RN001	RN002	RN003	RN004	RN005
RF 001					X
RF 002				X	
RF 003		X			
RF 004		X			
RF 005		X			
RF 006				X	X
RF 007		X			
RF 008		X		X	X
RF 009	X			X	X
RF 010	X				X
RF 011					

Fonte: Autoria própria.

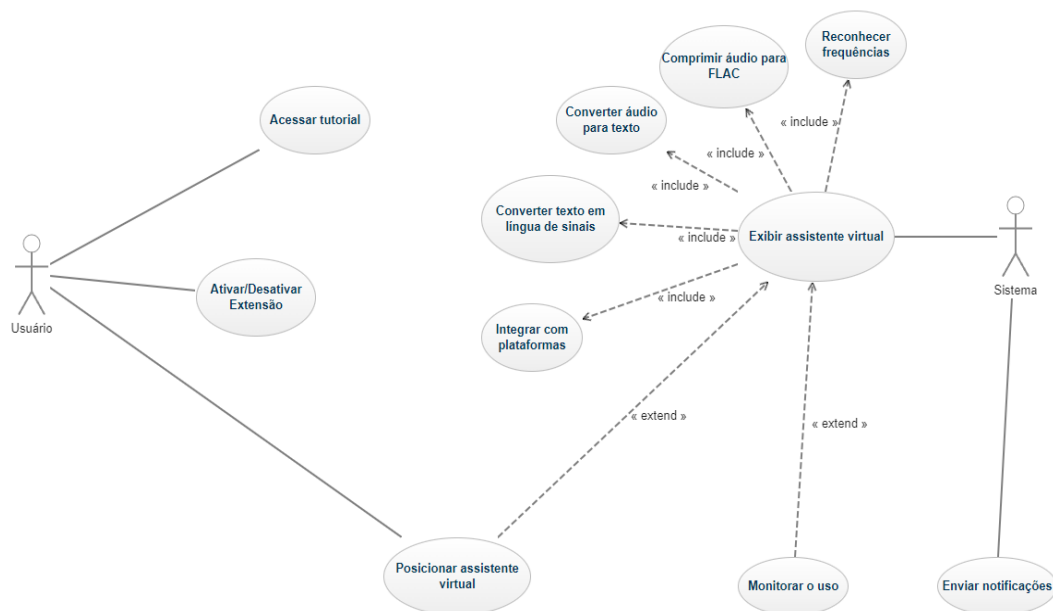
4.8 DIAGRAMA DE CASO DE USO

O diagrama de caso de uso é uma ferramenta utilizada na *Unified Modeling Language* (UML) para representar de forma visual os diferentes atores envolvidos em um sistema e as interações entre eles. É uma técnica muito útil para entender e definir os requisitos do sistema, permitindo que os desenvolvedores entendam como o usuário irá interagir com o sistema.

O diagrama de caso de uso é composto por atores (pessoas, sistemas ou outros elementos externos ao sistema) e casos de uso (ações realizadas pelos atores no sistema). O objetivo é ilustrar como o usuário interage com o sistema, e quais são as funcionalidades oferecidas por ele.

Por meio desse diagrama, na Figura 5, é possível identificar os fluxos de interação do usuário com o sistema, as restrições e as possibilidades de uso do *software*. Dessa forma, o diagrama de caso de uso é uma ferramenta importante para a fase de definição de requisitos do sistema.

Figura 5 - Diagrama de Caso de Uso



Fonte: Autoria própria.

4.9 DOCUMENTAÇÃO DE CASO DE USO

A documentação de caso de uso é uma técnica utilizada na engenharia de *software* para documentar o usuário com o sistema e os requisitos dele e está descrita no Quadro 8. O objetivo principal é fornecer informações específicas sobre como o sistema deve funcionar e como ele deve ser utilizado pelo usuário final. Para tal, foi utilizado um índice de caso de uso com o intuito de facilitar a visualização e possibilitar claro entendimento:

- UC 001 - Acessar o tutorial;
- UC 002 - Ativar/Desativar extensão;
- UC 003 - Posicionar assistente virtual;
- UC 004 - Enviar notificações;
- UC 005 - Exibir assistente virtual.

A documentação é composta por uma descrição detalhada do caso de uso, que inclui informações sobre os atores envolvidos, as pré-condições e pós-condições, os

fluxos básicos e alternativos, podendo ser utilizados como base para testes de aceitação e para validar se o sistema atende aos requisitos definidos.

Quadro 8 – Casos de Uso

Caso de Uso - Acessar o tutorial	
ID	UC 001
Descrição	Esse caso tem por objetivo redirecionar o usuário para um tutorial com o intuito de orientar sobre as funcionalidades da extensão.
Ator Primário	Usuário
Pré-condição	A extensão a ser instalada.
Cenário Principal	1 - A extensão é instalada; 2 - O sistema começa a exibir o tutorial; 3 - O usuário continua assistindo o tutorial até o sistema.
Pós-condição	Nenhuma
Cenário Alternativo	Usuário clica para pular o tutorial e ele se encerra.
Inclusão	Nenhuma
Extensão	Nenhuma
Caso de Uso – Ativar/Desativar extensão	
ID	UC 002
Descrição	Esse caso tem por objetivo verificar se o usuário vai manter a extensão ativa ou se irá desativá-la.
Ator Primário	Usuário
Pré-condição	A extensão estar instalada.
Cenário Principal	1 - A extensão está instalada; 2 - O usuário clica ícone da extensão; 3 - A extensão está ativada; 4 - O usuário desativa a extensão.
Pós-condição	O sistema desativa a extensão
Cenário Alternativo	1 - A extensão está desativada; 2 - O usuário efetua a ativação da extensão.
Inclusão	Nenhuma

Extensão	Nenhuma
Caso de Uso - Posicionar assistente virtual	
ID	UC 003
Descrição	Esse caso tem por objetivo escolher o posicionamento do assistente virtual na tela do usuário.
Ator Primário	Usuário
Pré-condição	A extensão estar exibindo o assistente virtual.
Cenário Principal	1 - A extensão está exibindo o assistente virtual; 2 - O usuário mantém pressionado o clique em cima do assistente virtual; 3 - O usuário reposiciona o assistente para o local desejado; 4 - O usuário solta o clique do botão.
Pós-condição	O assistente se mantém onde foi posicionado.
Cenário Alternativo	Nenhum
Inclusão	Nenhuma
Extensão	Nenhuma
Caso de Uso - Enviar notificações	
ID	UC 004
Descrição	Esse caso tem por objetivo o sistema enviar notificações de atualização e novidades sobre a extensão.
Ator Primário	Sistema
Pré-condição	Ter alterações no sistema da extensão.
Cenário Principal	1 - A extensão possui uma nova versão para atualização/novidade; 2 - O sistema sinaliza visualmente a existência de uma notificação; 3 - O usuário clica no ícone da extensão; 4 - O usuário lê a notificação.
Pós-condição	A sinalização visual de notificação não lida desaparece assim como a notificação em si.
Cenário Alternativo	Nenhum
Inclusão	Nenhuma

Extensão	Nenhuma
Caso de Uso - Exibir assistente virtual	
ID	UC 005
Descrição	Esse caso tem por objetivo exibir o assistente virtual na tela.
Ator Primário	Sistema
Pré-condição	Um vídeo estar sendo reproduzido em alguma plataforma audiovisual.
Cenário Principal	1 - O vídeo começa a ser reproduzido; 2 - O assistente virtual automaticamente aparece no canto inferior direito do conteúdo sendo exibido.
Pós-condição	O vídeo continua sendo reproduzido com a tradução concomitantemente.
Cenário Alternativo	A extensão está desativada e o assistente não é exibido automaticamente.
Inclusão	Reconhecer frequências; Comprimir áudio para FLAC; Converter áudio para texto; Converter texto em língua de sinais; Integrar com plataformas.
Extensão	Posicionar assistente virtual; Monitorar o uso.

Fonte: Autoria própria.

4.10 MATRIZ DE RASTREABILIDADE (REQUISITOS FUNCIONAIS X CASOS DE USO)

A matriz de rastreabilidade é uma ferramenta utilizada para correlacionar os requisitos de um sistema com os casos de uso que os implementam, como mostra o Quadro 9.

Essa ferramenta é útil para garantir que todos os requisitos definidos para o sistema sejam atendidos pelos casos de uso, e para identificar casos de uso que podem estar implementando mais de um requisito funcional.

Quadro 9 - Matriz de Rastreabilidade – RF X UC

	UC 001	UC 002	UC 003	UC 004	UC 005
RF 001	X	X			
RF 002		X	X		X
RF 003		X			X
RF 004					X
RF 005					X
RF 006			X		X
RF 007					X
RF 008			X		X
RF 009		X			X
RF 010		X		X	
RF 011		X		X	

Fonte: Autoria própria.

5 PROJETO CIDA

Neste capítulo, serão abordados os detalhes técnicos e as ferramentas utilizadas para o desenvolvimento da documentação e dos requisitos do projeto, assim como as bibliotecas e as tecnologias que serão necessárias para a realização da ferramenta.

5.1 FERRAMENTAS

As ferramentas utilizadas são recursos, tecnologias, *softwares*, *hardwares* ou métodos empregados em um projeto. Essas ferramentas têm como objetivo auxiliar na execução das tarefas, aumentar a eficiência e a qualidade do trabalho.

A escolha das ferramentas adequadas para cada fase do projeto e para as necessidades específicas das equipes envolvidas é fundamental para garantir um fluxo de trabalho mais ágil e eficiente, e para assegurar que os objetivos do projeto sejam alcançados de maneira satisfatória.

5.1.1 Python

Python (Figura 6) é uma linguagem de programação de alto nível e de propósito geral que é usada em diversas áreas, como desenvolvimento web, ciência de dados, Inteligência Artificial, automação de tarefas e muitas outras. Foi criada em 1991 por Guido van Rossum e é uma das linguagens mais populares do mundo, com uma grande comunidade de desenvolvedores e bibliotecas de código aberto (PYTHON, 2021).

Figura 6 – Python



Fonte: PYTHON, 2021, *online*.

Python é uma linguagem interpretada, o que significa que não é necessário compilar o código antes de executá-lo, permitindo uma maior agilidade no processo de desenvolvimento, além de facilitar a prototipagem e a experimentação.

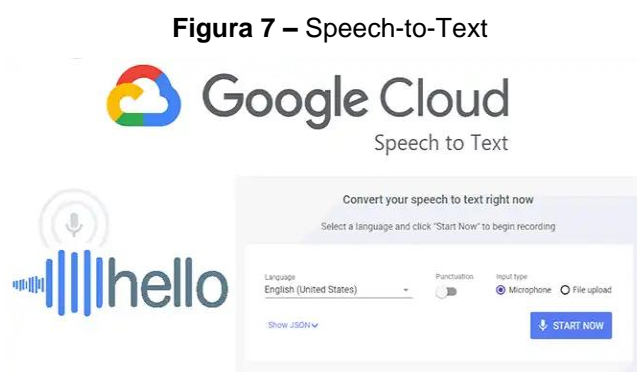
Uma das grandes vantagens dele é a sua simplicidade e clareza. A linguagem é fácil de aprender e de ler, com uma sintaxe intuitiva que se assemelha à linguagem natural, além de que possui uma ampla gama de bibliotecas disponíveis, o que simplifica muito o processo de desenvolvimento de projetos mais complexos.

Além disso, essa linguagem é multiplataforma, o que significa que pode ser executada em diferentes sistemas operacionais, como Windows, Mac e Linux, tornando-o uma ótima escolha para projetos que precisam ser executados em diferentes ambientes.

Essa linguagem será utilizada como ferramenta principal para desenvolvimento e manipulação das bibliotecas que possibilitarão a geração do projeto

5.1.2 Speech-To-Text

A API Speech-to-Text do Google Cloud Console (Figura 7) é uma tecnologia avançada de reconhecimento de voz que permite a conversão de áudio em texto. Com a ajuda dessa API, os desenvolvedores podem criar aplicativos que transcrevem a fala em tempo real, permitindo que os usuários falem com o aplicativo em vez de digitarem.



Fonte: SPEECHTOTEXT, 2022, *online*.

O reconhecimento de fala é realizado usando modelos de aprendizado de máquina avançados, que analisam a fala e a convertem em texto, sendo esses

modelos treinados em grandes conjuntos de dados de fala e texto para garantir alta precisão na transcrição, suportando uma variedade de idiomas e dialetos, e diferentes formatos de áudio, como MP3, WAV, FLAC e mais.

A API é usada em uma variedade de aplicativos, como serviços de transcrição de fala, aplicativos de reconhecimento de fala em tempo real, ferramentas de legenda automática para vídeos e muito mais. Com sua alta precisão e ampla variedade de recursos, o Speech-to-Text é uma das tecnologias mais avançadas disponíveis para reconhecimento de fala e transcrição de áudio em texto (SPEECHTOTEXT, 2022).

Sua principal utilidade na proposta de desenvolvimento da plataforma CIDA é para reconhecimento das frequências contidas no áudio do *streaming* de vídeo para posterior conversão em texto.

5.1.3 VLibras

O VLibras (Figura 8) é uma tecnologia desenvolvida pelo Ministério da Ciência, Tecnologia, Inovações e Comunicações do Brasil para auxiliar pessoas surdas e com deficiência auditiva a se comunicarem com mais facilidade. É uma plataforma que oferece tradução automática de texto em português para a Língua Brasileira de Sinais, permitindo que os usuários surdos ou com deficiência auditiva compreendam o conteúdo textual por meio de sinais (SOFTWARE PÚBLICO BRASILEIRO, *sd*).

Figura 8 – VLibras



Fonte: VLIBRAS, 2021, *online*.

A plataforma é composta por dois componentes principais: um *plugin* de acessibilidade, que pode ser instalado em sites para oferecer tradução de texto em tempo real para Libras, e um aplicativo móvel que oferece tradução de texto para Libras por meio da câmera do dispositivo móvel, também oferecendo um dicionário

online com mais de 6.000 sinais em Libras, um curso de Libras para iniciantes e um serviço de videoconferência com intérpretes de Libras.

A tecnologia VLibras tem sido amplamente utilizada no Brasil para tornar a comunicação mais acessível para pessoas surdas e com deficiência auditiva, com sua plataforma abrangente e recursos avançados, ela ajuda a quebrar barreiras de comunicação e promover inclusão em diversos contextos, desde o acesso a conteúdo *online* até a comunicação em tempo real.

Essa será a principal biblioteca utilizada na aplicação devido ao fato de ser uma ferramenta *open source* e disponibilizar um modelo 3D para representar graficamente o texto traduzido em linguagens de sinais.

5.1.4 FFmpeg

A ferramenta FFmpeg (Figura 9) é uma coleção de *software* livre para manipulação de áudio e vídeo, com recursos para decodificação, codificação e transcodificação de formatos de arquivo, e é um conjunto de ferramentas de linha de comando e bibliotecas de programação que podem ser utilizadas em diversas plataformas, incluindo Linux, Windows e macOS (TECHTUDO, *sd*).

Figura 9 – FFmpeg



Fonte: FFMPEG, *sd*, *online*.

Com essa tecnologia, os desenvolvedores podem criar aplicativos para processar e manipular arquivos de áudio e vídeo em diferentes formatos, como *MPEG Layer 4* (MP4), *Audio Video Interleaved* (AVI), *Matroska Video* (MKV), WAV, MP3 e muitos outros. A biblioteca suporta uma ampla variedade de codecs, permitindo que os usuários convertam arquivos de um formato para outro.

No CIDA, o principal uso será para comprimir o áudio extraído em texto de forma a evitar perdas de informações e manter a qualidade do produto gerado, visando proporcionar uma tradução completa e livre de erros semânticos e sintáticos.

5.1.5 Visual Studio Code

A IDE sugerida para manusear de forma prática o desenvolvimento da aplicação, o Visual Studio Code (Figura 10) é um editor de código-fonte gratuito e de plataforma cruzada, desenvolvido pela Microsoft. É uma ferramenta de desenvolvimento altamente flexível e personalizável, que oferece recursos avançados para ajudar os desenvolvedores a escrever, depurar e testar código de forma eficiente (REMESSA ONLINE, 2021).

Figura 10 – Visual Studio Code



Fonte: VSCODE, 2021, *online*.

5.1.6 Miro

Miro (Figura 11) é uma ferramenta de colaboração visual online que permite que indivíduos e equipes colaborem, criem, compartilhem e visualizem ideias de forma digital. É uma plataforma de quadro branco virtual que ajuda as equipes a colaborar e organizar suas ideias de forma mais eficiente, amplamente utilizado por equipes de negócios, de *design* e de desenvolvimento para colaboração e *brainstorming* em tempo real, bem como para gerenciamento de projetos e processos de criação de produtos (GOMES, 2022).

Figura 11 – Miro

Fonte: MIRO, *sd, online*.

Foi utilizado principalmente para representação gráfica de elementos de organização e visuais, como a matriz SWOT, por exemplo.

5.1.7 GenMyModel

O GenMyModel (Figura 12) é uma ferramenta *online* para modelagem e design de diagramas de classe UML, que permite que os usuários criem, visualizem e colaborem em modelos de dados e processos de forma intuitiva e fácil (GENMYMODEL, 2023).

Figura 12 – GenMyModel

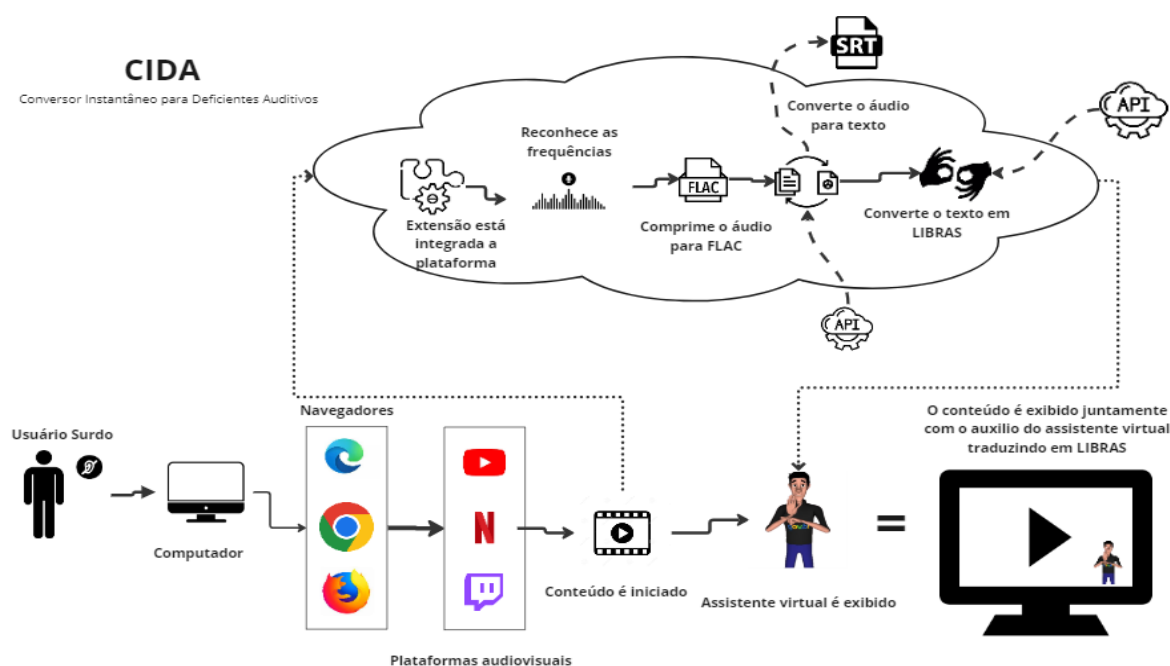
Fonte: GENMYMODEL, 2023, *online*

Utilizando-o, foi possível o desenvolvimento das principais representações dos requisitos da aplicação por meio dos diagramas de caso de uso e BPMN, facilitando o entendimento do fluxo da ferramenta trabalhada.

5.2 FUNCIONALIDADES DO SISTEMA

A principal funcionalidade do sistema envolve o processo de exibição e utilização do modelo 3D do assistente virtual, para tal, na Figura 13 tem-se uma visualização de como será o fluxo tanto pela ótica de um usuário da plataforma, quanto a demonstração das ferramentas e o processo que é executado em *background*.

Figura 13 – Fluxo CIDA



Fonte: Autoria própria.

Como se pode observar, o fluxo se inicia com o acesso a um navegador no qual a ferramenta foi instalada e, após isso, a reprodução começa com o acesso às plataformas onde a ferramenta possui integração. Dado esse momento em que o conteúdo é reproduzido, se inicia os procedimentos realizados no *back-end* da aplicação que envolve a utilização de equações para reconhecer corretamente as frequências de áudio e isolá-las para que, assim, seja possível comprimi-las para um formato sem perda de dados (no fluxo foi utilizado a compressão para o formato FLAC).

Feito isso, o próximo passo é enviar o conteúdo de áudio limpo para um serviço externo responsável por convertê-lo para o formato de texto que necessário para prosseguir. Para tal, a recomendação é a utilização da plataforma Speech To Text (STT) para conversão baseada em aprendizado de máquina, licenciada pela Google e disponível no ambiente *Google Cloud Platform* (GCP), que oferece diversos recursos e serviços que variam desde ferramentas de análise de dados até o uso de Inteligência Artificial, como é o caso do CIDA.

Com o arquivo retornado na extensão *SubRip Subtitle* (SRT), o passo final é disponibilizar a legenda gerada para outra ferramenta externa capaz de ler, interpretar e renderizar o modelo de assistente virtual que será utilizado para transcrever em linguagem de sinais o conteúdo enviado ao serviço. Como o objetivo principal da extensão CIDA é proporcionar um meio simples, intuitivo e acessível para quem opte por utilizá-la, a prioridade de escolha de bibliotecas externas foi sua disponibilidade sem grandes custos (preferencialmente *open source*). Mediante isso, a solução identificada e recomendada é a utilização da API VLibras para esse fim.

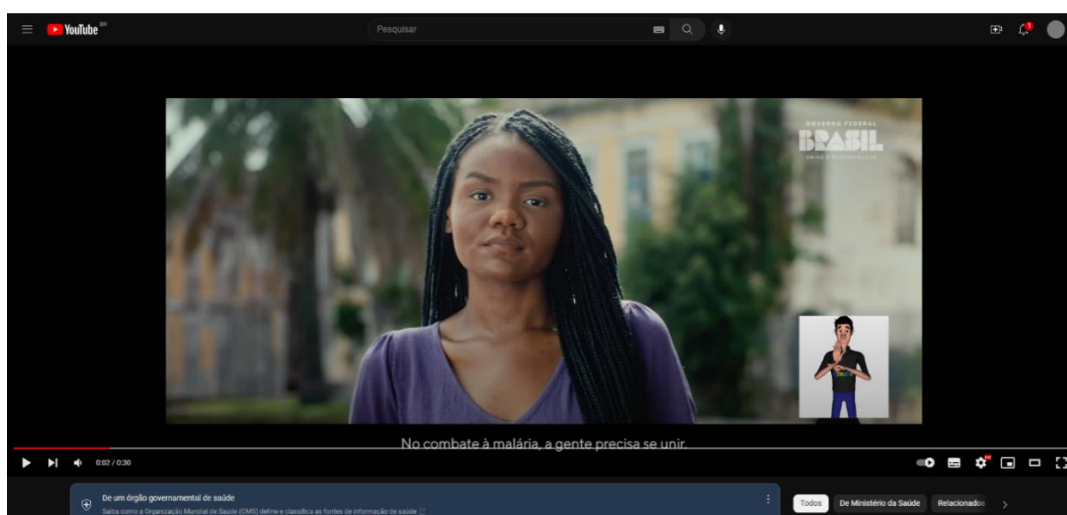
5.3 TELAS DO SISTEMA

Telas do sistema são representações gráficas das interfaces de usuário que fazem parte de um sistema. Essas telas variam de acordo com o tipo de sistema e o objetivo do projeto, mas em geral, incluem elementos como menus, botões, campos de entrada de dados, caixas de diálogo e outros recursos de interação do usuário.

A criação de telas do sistema é uma etapa importante no desenvolvimento de um *software*, pois permite que os desenvolvedores visualizem como o sistema será utilizado pelos usuários e testem a usabilidade e eficiência das interfaces. Além disso, as telas do sistema também são importantes para a documentação do projeto, permitindo que os desenvolvedores e usuários finais compreendam como o sistema funciona e como interajam com ele.

Na Figura 14, é possível observar o modelo de visualização da ferramenta sendo utilizada em um vídeo do Youtube.

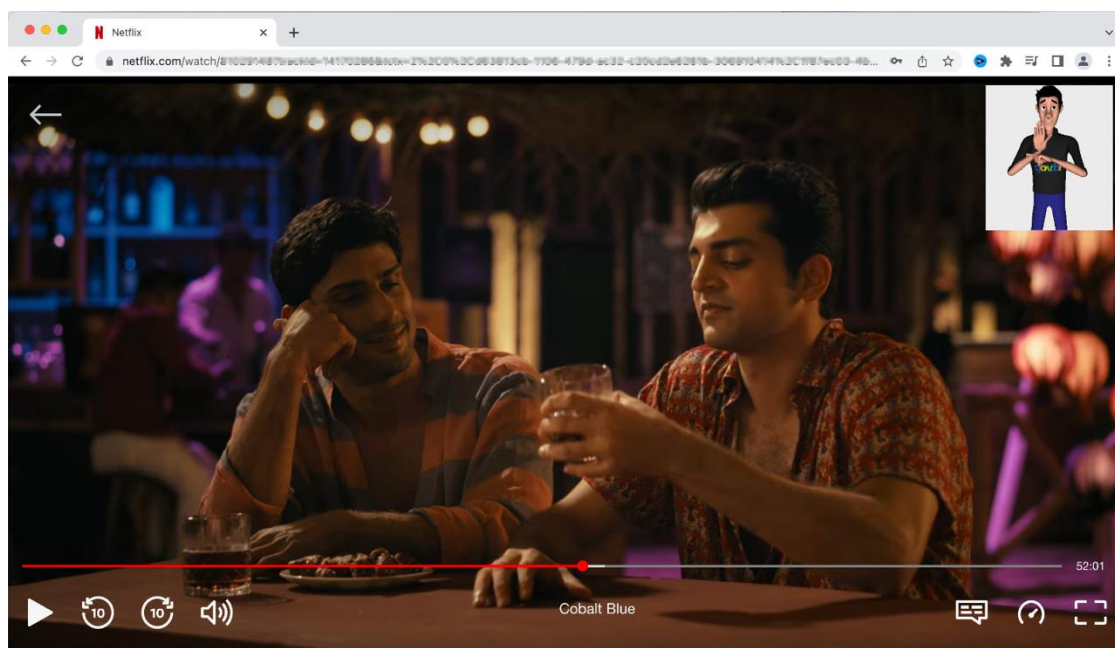
Figura 14 – Tela Youtube



Fonte: Autoria própria.

Já a Figura 15 apresenta a ferramenta sendo utilizada em *streaming* de vídeo online (Netflix), traduzindo uma conversa encenada com dois atores de um seriado para linguagem de sinais em tempo real.

Figura 15 – Tela Netflix

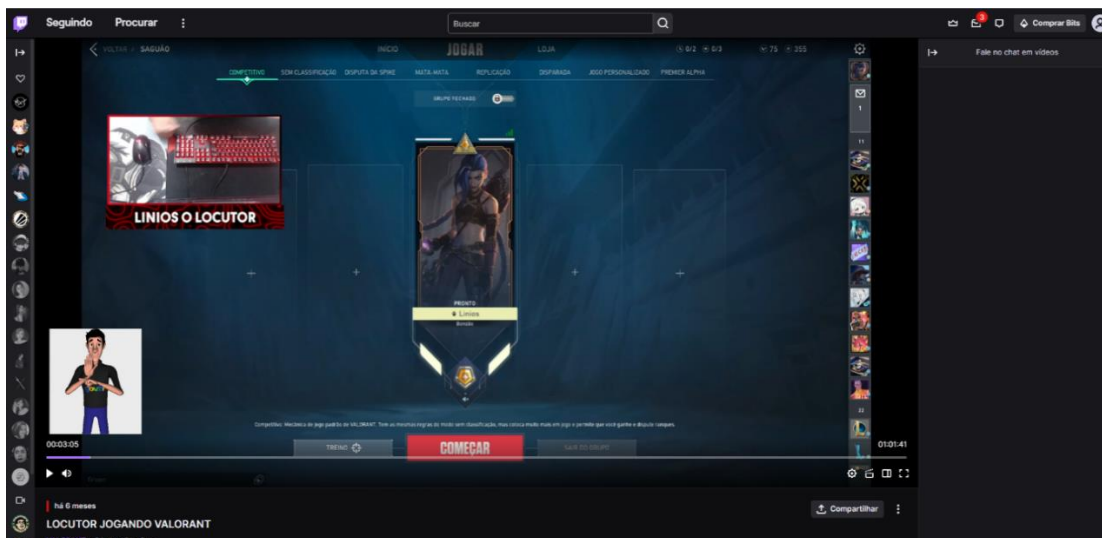


Fonte: Autoria própria.

Na Figura 16 é possível notar o *plugin* em funcionamento em um *streaming* de conteúdo dinâmico da plataforma Twitch, já utilizando a funcionalidade responsável

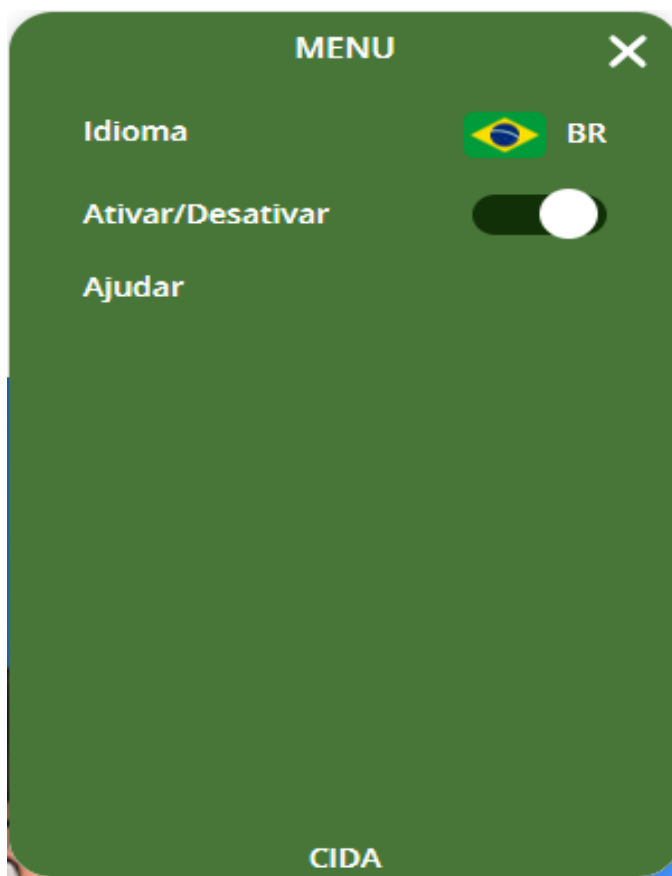
por alterar o posicionamento do assistente virtual para diferentes posições da tela do vídeo.

Figura 16 – Tela Twitch



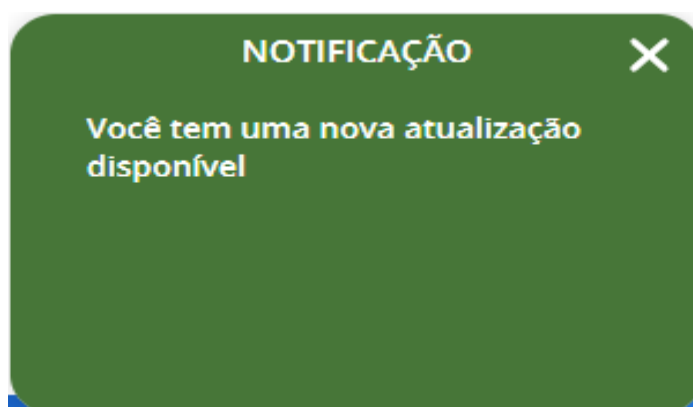
Fonte: Autoria própria.

Na Figura 17 pode ser visualizado o menu responsável por permitir a ativação e desativação da extensão, conforme necessidade do usuário, assim como um botão para posterior exibição do tutorial de funcionamento da ferramenta.

Figura 17 – Tela Menu

Fonte: Autoria própria.

Na Figura 18 tem a visualização do bloco no qual estarão as notificações de atualizações e melhorias da ferramenta para orientação e conhecimento do usuário final, de forma a sempre mantê-lo atualizado do progresso de correções da extensão.

Figura 18 – Tela Notificação

Fonte: Autoria própria.

5.4 BENEFÍCIOS

Dentre os benefícios que esta solução proporciona, tem-se a contribuição com os Objetivos para o Desenvolvimento Sustentável (ODS), que são objetivos ambiciosos e interconectados que abordam os principais desafios de desenvolvimento enfrentados por pessoas no Brasil e no mundo, criados pela Organização das Nações Unidas. Entre os principais indicadores brasileiros, tem-se a erradicação da pobreza, igualdade de gênero, trabalho decente e crescimento econômico, além, é claro, do ODS 10, que se trata da redução das desigualdades, tema esse, que gostaríamos de ressaltar. Esse objetivo tem como uma de suas metas o empoderamento e a promoção da inclusão social, econômica e política de todos, independentemente da idade, gênero, deficiência, raça, etnia, origem, religião, condição econômica ou outra. Inclusive, o seu indicador de sucesso é pautado na proporção de pessoas que vivem abaixo de 50% da média da renda, por sexo, idade e pessoas com deficiência. Quando pensamos então em como promover essa inclusão socioeconômica, principalmente ao falar em pessoas com deficiência, a primeira palavra que vem à mente, é claro, é acessibilidade, seja ela arquitetônica, atitudinal, comunicacional, ou digital, todas as suas formas têm papel de destaque no combate à desigualdade (ESTRATEGIA ODS, *sd*).

Além dos requisitos a se cumprir referente ao desenvolvimento sustentável, pode-se citar também as boas práticas da *Environment, Social & Governance* (ESG), uma abordagem para avaliar até que ponto uma corporação trabalha em prol de objetivos sociais que vão além do papel de uma corporação para maximizar os lucros em nome dos acionistas da empresa. O eixo ambiental da sigla *Environment* diz respeito às políticas em prol do meio ambiente e como a empresa se mobiliza interna e externamente em relação a esse assunto, que é bastante amplo. Já o eixo Social da ESG refere-se às ações que promovem diversidade e inclusão social, envolvendo pautas como o respeito e apoio aos Direitos Humanos, direitos trabalhistas, equidade de gênero, ambiente de trabalho saudável, relação com a comunidade local, dentre outras. O eixo da governança (*Governance*) encontra eco nas práticas transparentes, éticas e responsáveis na administração de uma organização. Trata-se de um elemento competitivo relevante, uma vez que as boas práticas de governança

preveem estratégias de gestão mais eficientes e confiáveis, algo que pesa na decisão de investidores, por exemplo (XP EDUCAÇÃO, 2022).

Considerando esses indicadores, a principal proposta da ferramenta é proporcionar a inclusão de pessoas que vivem à margem da sociedade devido a uma deficiência que lhes foi imposta muitas vezes desde o nascimento. Para tal, aumentar a praticidade e possibilitar a visualização de conteúdos multimídia aumentará as oportunidades de acesso a conhecimento, cultura e entretenimento para usuários deficientes auditivos.

5.5 POSSÍVEIS LIMITAÇÕES E DESAFIOS

Frente a proposta de desenvolvimento do CIDA, verificou-se alguns desafios a superar que dificulta a criação e sustentação da ferramenta, como:

- **Infraestrutura:** a infraestrutura necessária para manter a plataforma em execução, já que a ferramenta exige a comunicação em tempo real com diversas bibliotecas externas, assim como espera uma resposta adequada para que possa prosseguir com o fluxo e para que assim, possa prover o resultado esperado. Uma estrutura para esse fim se torna extremamente robusta em âmbito de servidores, recursos em nuvem e provedores de internet, gerando custos elevados para se manter a aplicação em execução.
- **Internet:** uma conexão estável com a internet é fundamental para que todos os processos funcionem de forma correta.
- **Pontos de Oscilação:** a API de reconhecimento de frequência de fala tem dificuldade em capturar os dados do áudio corretamente quando há pontos de oscilação, como falas muito rápidas ou baixas, que podem ser identificados como ruídos externos ou de ambiente.
- **Disponibilidade:** disponibilidade do VLibras foi um problema que dificultou a progressão do estudo da criação da ferramenta, já que a API atualmente se encontra indisponível e, infelizmente, a ferramenta não possui uma documentação elaborada acerca de seu uso e aplicação para simulação em desenvolvimento local.

CONSIDERAÇÕES FINAIS

A falta de acessibilidade para deficientes auditivos é um problema recorrente na sociedade, especialmente quando se trata de conteúdos audiovisuais. Mesmo com a presença de legendas, muitas vezes essas não são suficientes para permitir a compreensão completa do conteúdo. Além disso, muitos deficientes auditivos são analfabetos ou semianalfabetos, o que torna a leitura de legendas inviável.

A proposta do Projeto CIDA mostra-se como uma solução inovadora e acessível para a inclusão digital de pessoas surdas, utilizando tecnologias modernas para a criação de um assistente virtual, em paralelo a mecanismos de reconhecimento de voz, processamento de linguagem natural e tradução de áudio para texto e, por fim, para língua de sinais, dessa forma, será possível oferecer um recurso de acessibilidade mais eficaz e abrangente para deficientes auditivos, permitindo-lhes consumir conteúdos audiovisuais de forma mais acessível e eficiente.

Como trabalhos futuros, sugere-se a realização de testes com usuários surdos para avaliar a efetividade e usabilidade do CIDA, além de possíveis melhorias na interface do sistema e a incorporação de novas funcionalidades e linguagens de sinais que não abrangem apenas a Língua Brasileira de Sinais, como a *American Sign Language* e a *Língua Gestual Portuguesa*, por exemplo. Já quanto a novas implementações, tem-se a possibilidade de customização do assistente virtual, de forma a permitir o usuário adaptar o modelo 3D, garantindo assim, a representatividade na ferramenta. Outra proposta de melhoria seria a oportunidade de tornar a extensão acessível em grande parte dos navegadores disponíveis no mercado e garantir que a tradução para língua de sinais seja realizada independente da língua de origem no qual o áudio do *streaming* ou vídeo é disponibilizado. Vale também considerar a expansão da plataforma para incluir outros ambientes de comunicação, como reuniões, palestras e eventos.

Por fim, a realização deste projeto possibilitou a aplicação dos conhecimentos teóricos e práticos adquiridos ao longo do curso de Análise e Desenvolvimento de Sistemas, como técnicas de especificação de requisitos e modelagem de processos de negócio, além do desenvolvimento de habilidades e competências em áreas como a engenharia de *software*, processamento de linguagem natural e interação humano-

computador. A partir das experiências vivenciadas neste projeto, fica evidente a importância de se investir em soluções tecnológicas e, por sua vez, no uso da tecnologia como ferramenta para a inclusão social e melhoria da qualidade de vida das pessoas com deficiência auditiva, promovendo uma sociedade cada vez mais acessível para todos.

REFERÊNCIAS

ANDRADE, M. **LanCap - Tradutor de vídeos para Libras**. 04/042019. Disponível em: <<https://github.com/mateusandrade98/lancap>>. Acesso em: 20.jun.2022.

ASSISTENTES. **VLibras**. 24/05/2023. Disponível em: <<https://www.gov.br/governodigital/pt-br/vlibras>>. Acesso em: 06.mai.2023.

BIBLIO FORA DA CAIXA. **Reconhecimento de Voz: a tecnologia por trás dos assistentes virtuais**. 19/02/2023. Disponível em: <<https://biblioforadacaixa.com.br/index.php/2023/02/19/reconhecimento-de-voz-a-tecnologia-por-tras-dos-assistentes-virtuais/>>. Acesso em: 15.abr.2023.

BRASIL. **Dispõe sobre a Língua Brasileira de Sinais - Libras e dá outras providências**. 24/04/2002. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/2002/L10436.htm>. Acesso em: 15.abr.2023.

BRASIL. **Estabelece normas gerais e critérios básicos para a promoção da acessibilidade das pessoas portadoras de deficiência ou com mobilidade reduzida, e dá outras providências**. 19/12/2000. Disponível em: <https://www.planalto.gov.br/ccivil_03/leis/l10098.htm>. Acesso em: 15.abr.2023.

BRASIL. **Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência)**. 06/07/2015. Disponível em: <https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2015/Lei/L13146.htm>. Acesso em: 30.abr.2023.

CRISTIANO, A. **O que é Libras**. 17/05/2017. Disponível em: <<https://www.libras.com.br/o-que-e-libras>>. Acesso em: 07.abr.2023

CYRINO, L. **Aprendizado de Máquina (Machine Learning)**. 17/11/2021. Disponível em: <<https://www.manutencaoemfoco.com.br/aprendizado-de-maquina-machine-learning/>>. Acesso em: 08.mai.2023.

DEEP LEARNING BOOK. **Capítulo 77 – Modelo BERT para Processamento de Linguagem Natural**. *sd*. Disponível em: <<https://www.deeplearningbook.com.br/modelo-bert-para-processamento-de-linguagem-natural/>>. Acesso em: 15.abr.2023.

ESTADÃO. **Raio-X da surdez no Brasil**. 09/2019. Disponível em: <https://www.estadao.com.br/blogs/blog/wp-content/uploads/sites/189/2019/09/RaioXSurdez_InstitutoLocomotiva_30setembro2019_blogVencerLimites.pdf>. Acesso em: 30.abr.2023.

ESTRATEGIA ODS. **ODS10**. *sd*. Disponível em: <<https://www.estrategiaods.org.br/os-ods/ods10/>>. Acesso em: 30.abr.2023.

FFMPEG. **FFmpeg Project**. *sd*. Disponível em: <<https://trac.ffmpeg.org/>>. Acesso em: 08.mai.2023.

FORBES. **Menos de 1% dos sites brasileiros são considerados acessíveis, diz pesquisa**. 29/07/2021. Disponível em: <<https://forbes.com.br/forbesesg/2021/07/mentos-de-1-dos-sites-brasileiros-sao-considerados-acessiveis-diz-pesquisa/>>. Acesso em: 09.mai.2023.

GALA, A. S. **Leis de acessibilidade no Brasil: tudo o que você precisa saber**. 15/12/2022. Disponível em: <<https://www.handtalk.me/br/blog/leis-de-acessibilidade/>>. Acesso em: 30.abr.2023.

GENMYMODEL. **Why GenMyModel is different**. *sd*. Disponível em: <<https://www.genmymodel.com/>>. Acesso em: 03.mai.2023.

GÉRON, A. **HANDS-ON MACHINE LEARNING WITH SCIKIT-LEARN, KERAS, AND TENSORFLOW**. 2019.

GOMES, M. **Miro: ferramenta gratuita para trabalhar colaborativamente**. 16/03/2022. Disponível em: <<https://www.alura.com.br/artigos/miro-ferramenta-gratuita-trabalhar-colaborativamente>>. Acesso em: 07.mai.2023.

HANDTALK. **Aplicativo HandTalk**. *sd*. Disponível em: <<https://www.handtalk.me/br/aplicativo/>>. Acesso em: 06.mai.2023.

HAYKIN, S. **REDES NEURAIIS: PRINCÍPIOS E PRÁTICA**. 2001.

HONORA, M.; FRIZANCO, M.L.E. **Livro Ilustrado de Língua Brasileira de Sinais**. 2009.

IBM. **O que são redes neurais?** *sd*. Disponível em: <<https://www.ibm.com/br-pt/topics/neural-networks>>. Acesso em: 08.mai.2023.

INBOT. **Assistentes Virtuais**. 12/01/2023. Disponível em: <<https://www.inbot.com.br/assistentes-virtuais/assistentes-virtuais/>>. Acesso em: 08.mai.2023.

LANGSTON, J. **Microsoft e Nuance unem forças para ajudar médicos a manterem foco nos pacientes**. 07/01/2020. Disponível em: <<https://news.microsoft.com/pt-br/features/microsoft-e-nuance-unem-forcas-para-ajudar-medicos-a-manterem-foco-nos-pacientes/>>. Acesso em: 15.abr.2023.

MICROSOFT, **Microsoft lança recurso de visualização de linguagem de sinais para reuniões do Teams**. 18/11/2022. Disponível em: <<https://news.microsoft.com/pt-br/microsoft-lanca-recurso-de-visualizacao-de-linguagem-de-sinais-para-reunioes-do-teams/>>. Acesso em: 15.abr.2023.

MIRO. **THE ONLINE COLLABORATIVE WHITEBOARD PLATFORM TO BRING TEAMS TOGETHER, ANYTIME, ANYWHERE.** *sd.* Disponível em: <<https://www.gend.co/miro/>>. Acesso em: 01.jun.2023.

PYTHON. **As vantagens de aprender Python.** 16/11/2021 Disponível em: <<https://medium.com/data-hackers/quais-s%C3%A3o-as-vantagens-de-aprender-python-dd3fbba02c3c>>. Acesso em: 08.mai.2023.

REMESSA ONLINE. **Visual Studio Code: confira as principais funções da ferramenta.** 26/10/2021. Disponível em: <<https://www.remissaonline.com.br/blog/visual-studio-code-confira-as-principais-funcoes-da-ferramenta/>>. Acesso em: 06.mai.2023.

SAFRAN. **SAFRAN IDENTITY & SECURITY apresenta soluções biométricas contactless no CIAB Febraban.** *sd.* Disponível em: <<https://www.abrid.org.br/safran-identity-security-apresenta-solucoes-biometricas-contactless-no-ciab-febraban-2/>>. Acesso em: 15.abr.2023.

SANCHEZ, C. **Assistentes Virtuais & Linguagem Natural: como implementar?** *sd.* Disponível em: <<https://botmaker.com/pt/publicacoes/assistentes-virtuais-linguagem-natural/>>. Acesso em: 15.abr.2023.

SANTOS, L. **Saiba tudo sobre arquivos FLAC.** *sd.* Disponível em: <<https://recoverit.wondershare.com.br/audio-repair/what-is-flac.html>>. Acesso em: 15.abr.2023.

SOFTWARE PÚBLICO BRASILEIRO. **Suite VLibras.** *sd.* Disponível em: <<https://softwarepublico.gov.br/social/suite-vlibras>>. Acesso em: 15.abr.2023.

SPEECHTOTEXT. **Google Releases New Cloud Speech-to-Text API.** 05/11/2022. Disponível em: <<https://blog.systoolsgroup.com/google-releases-new-cloud-speech-to-text/>>. Acesso em: 08.mai.2023.

TECH.EU. **Meta acquires Audio Analytic.** 09/11/2022. Disponível em: <<https://tech.eu/2022/11/09/meta-acquires-audio-analytic>>. Acesso em: 29.mai.2023.

TECHTUDO. **FFmpeg: saiba tudo sobre a ferramenta de conversão e edição de vídeos.** *sd.* Disponível em: <<https://www.techtudo.com.br/tudo-sobre/ffmpeg/>>. Acesso em: 08.mai.2023.

UNITED NATIONS. **Standard Rules on the Equalization of Opportunities for Persons with Disabilities.** 24/06/2003. Disponível em: <<https://www.un.org/esa/socdev/enable/rights/contrib-wfd.htm>>. Acesso em: 30.abr.2023.

VLIBRAS. **Conheça o VLibras e aprenda como utilizá-lo.** 06/08/2021. Disponível em: <<https://www.uftm.edu.br/home/comunicados/96-geral/3405-conheca-o-vlibras-e-aprenda-como-utiliza-lo>>. Acesso em: 08.mai.2023.

VOCALCOM. **Reconhecimento de voz e síntese vocal em call center: por que utilizar?** *sd*. Disponível em: <<https://www.vocalcom.com/pt/blog/reconhecimento-voz-sintese-vocal-call-center/>>. Acesso em: 30.abr.2023.

VOCÊ S/A. **Hand Talk, a startup que ajuda na comunicação com surdos.** 15/10/2021. Disponível em: <<https://vocesa.abril.com.br/empreendedorismo/hand-talk-a-startup-que-ajuda-na-comunicacao-com-surdos/>>. Acesso em: 07.abr.2023.

VSCODE. **Abrir VS Code via terminal no MacOS.** 28/06/2021. Disponível em: <<https://coffops.com/abrir-vs-code-via-terminal-no-macos/>>. Acesso em: 08.mai.2023.

XP EDUCAÇÃO. **5 práticas ESG para implementar na sua empresa e conquistar o selo.** 15/12/2022. Disponível em: <<https://blog.xpeducacao.com.br/praticas-esg/>>. Acesso em: 30.abr.2023.