

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE ARQUITETURA, ARTES, COMUNICAÇÃO E DESIGN
PROGRAMA DE PÓS-GRADUAÇÃO EM MÍDIA E TECNOLOGIA**

FELIPE RODOLFO DOMINGUES DA SILVA

**ISA — INTERFACE DE SUPORTE À ACESSIBILIDADE: UM PROTÓTIPO
DE CLOSED CAPTION ACESSÍVEL**

BAURU

2022

FELIPE RODOLFO DOMINGUES DA SILVA

**ISA — INTERFACE DE SUPORTE À ACESSIBILIDADE: UM PROTÓTIPO
DE CLOSED CAPTION ACESSÍVEL**

Relatório Técnico da tese de Mestrado apresentado ao Programa de Pós-Graduação em Mídia e Tecnologia, da Faculdade de Arquitetura, Artes e Comunicação – FAAC, Universidade “Júlio de Mesquita Filho” – UNESP, como requisito parcial para obtenção do título de Mestre em Mídia e Tecnologia sob a orientação do Prof. Dr. Francisco Machado Filho

**BAURU
2022**

S586i Silva, Felipe Rodolfo Domingues da
ISA —Interface de Suporte à Acessibilidade : um
protótipo de closed caption acessível / Felipe
Rodolfo Domingues da Silva. -- Bauru, 2022
66 p. : tabs., fotos

Dissertação (mestrado) - Universidade Estadual
Paulista (Unesp), Faculdade de Arquitetura, Artes,
Comunicação e Design, Bauru
Orientador: Francisco Machado Filho

1. Acessibilidade. 2. Tecnologia. 3. Surdez. 4.
Legendagem. 5. Python. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp.
Biblioteca da Faculdade de Arquitetura, Artes, Comunicação e Design,
Bauru. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

FELIPE RODOLFO DOMINGUES DA SILVA

**ISA — INTERFACE DE SUPORTE À ACESSIBILIDADE: UM PROTÓTIPO
DE CLOSED CAPTION ACESSÍVEL**

Banca examinadora:

Prof. Dr. Francisco Machado Filho

Presidente / Orientador / FAAC – UNESP Bauru

Prof. Dra. Celia Maria Retz Godoy Santos

Docente / FAAC – UNESP Bauru

Prof. Dr. Vitor Pachioni Brumatti

Docente – UNISAGRADO Bauru

Resultado: _____

Bauru, _____ / _____ / _____

Dedico essa pesquisa à minha amada esposa Isadora, a qual além de ser sempre a minha fonte de inspiração, é também meu alicerce e meu porto seguro.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ter me dado a oportunidade e me ajudado a chegar até aqui, sei que antes de tudo foi Ele quem planejou isso para mim.

Em segundo agradeço minha família, em especial meu pai José (*in memoriam*), e minha irmã Nara (*in memoriam*), que foram meus principais incentivadores e educadores. Com certeza, um dos meus maiores desejos era poder dar-lhes um abraço agora e poder dizer: “Eu consegui!”.

Agradeço também, à minha querida professora de ensino fundamental Angela Assis (*in memoriam*), que através de poucas palavras, mas de atos grandiosos e graciosos, me fez acreditar ser na simplicidade da vida onde se encontram as reais essências do ser humano. Infelizmente, foi mais uma vítima do vírus, mas quero que saiba, onde quer que você esteja, que sim, você foi muito importante na minha caminhada pessoal, profissional e acadêmica.

Aos meus queridos amigos meu carinhoso agradecimento por compartilharem comigo todos os momentos e batalhas até aqui, com certeza vocês são os melhores do mundo e fazem jus à expressão: “Amigos são irmãos e anjos que a vida nos dá para nos proteger e dividir as dores da peregrinação”.

Não posso esquecer de agradecer meu grande amigo, primeiro ex orientador Fábio Cardoso, que com seus conhecimentos técnicos me instruiu e auxiliou para eu conseguir chegar ao melhor resultado possível nessa pesquisa.

E finalizo, agradecendo ao meu orientador Prof. Dr. Francisco de Machado Filho, por ter a coragem em me aceitar como orientando mesmo eu não o tendo indicado como orientador e não ter muita afinidade com a área acadêmica, tendo apenas pelo brilho nos olhos e vontade de fazer acontecer. Por todos os desafios que passamos nesse período, ele foi fundamental para que tudo desse certo, uma das fontes da minha resiliência. Muito obrigado.

"Algumas pessoas enxergam a inovação como mudança, mas nós nunca vimos dessa forma. É sobre melhorar as coisas"

Tim Cook, ex-CEO da Apple

SILVA, F. R. D. da, **ISA – Interface de suporte à acessibilidade: um protótipo de *closed caption* acessível**. Programa de Pós- Graduação em Mídia e Tecnologia, da Faculdade de Arquitetura, Artes e Comunicação – FAAC, Universidade “Júlio de Mesquita Filho” — UNESP, sob a orientação do Prof. Dr. Francisco Machado Filho, 2022.

RESUMO

O objetivo deste estudo - sobre o uso do *Closed Caption* ou de tecnologias assistivas voltadas à população surda e ensurdecida - é realizar uma análise de desempenho desta e desenvolver um protótipo de Software de legendagem compatível com computadores pessoais e produtos audiovisuais semelhante a este, porém com melhor custo-benefício utilizando-se de tecnologias de *Speech-to-text* e módulos de inteligência artificial para aprimorar sua acessibilidade. Traz ainda a problematização do uso do *Closed Caption*, em relação ao custo benéfico, acessibilidade, desempenho e assertividade. Para isso, foi desenvolvido um levantamento bibliográfico sobre as temáticas de acessibilidade para surdos e ensurdecidos, com consultas a artigos acadêmicos e pesquisas de campo e estudos e análises de desempenho de softwares, de transcrição de áudio em texto, presentes no mercado. A biblioteca de transcrição de texto empregada foi a *Speech Recognition*, amplamente empregada em outros países como soluções de acessibilidade e transcrição de texto. O principal resultado foi a criação do protótipo similar a este acrescido de melhorias como a diminuição do atraso, a assertividade na fala junto ao receptor, a fácil implementação e uso e a viabilização econômica de baixo custo, podemos atender TVs universitárias, ONGs, vídeos e produções audiovisuais caseiras entre outras.

PALAVRAS-CHAVE

Acessibilidade. Tecnologia. Surdez. Legendagem. *Closed Caption*. Python. Inteligência Artificial

ABSTRACT

The objective of this study - on the use of Closed Caption or assistive technologies aimed at the deaf and deaf population - is to perform a performance analysis of this and develop a prototype of a subtitling software compatible with personal computers and audiovisual products similar to this one, but with more cost-effective using Speech-to-text technologies and artificial intelligence modules to improve its accessibility. It also brings up the problematization of the use of Closed Caption, in relation to the beneficial cost, accessibility, performance and assertiveness. For this, a bibliographic survey was developed on the themes of accessibility for the deaf and deaf, with consultations to academic articles and field research and studies and analysis of performance of software, from audio transcription to text, present in the market. The text transcription library used was Speech Recognition, widely used in other countries as accessibility and text transcription solutions. The main result was the creation of a prototype similar to this one, plus improvements such as reduced delay, assertiveness in speech to the receiver, easy implementation and use and low-cost economic feasibility, we can serve university TVs, NGOs, videos and homemade audiovisual productions among others.

KEYWORDS

Accessibility. Technology. Deafness. Subtitling. Closed Caption. Python. Artificial intelligence

LISTA DE FIGURAS

Figura 1 - Cecil M. Hepworth e os primeiros intertítulos oralizados do cinema mundial.....	10
Figura 2 - Os intertítulos utilizados por Robert Paul em ‘Scrooge, or, Marley’s Ghost’ (1901) e a utilização de Hepworth ‘Alice no País das Maravilhas’ (1903).....	10
Figura 3 - Edwin S. Porter e seus ‘textos verbais’ em ‘College Chums’ (1907).....	11
Figura 4 - Dois frames de “The Chamber Mystery” (1920) utilizando a técnica de balões de palavras.....	13
Figura 5 - Documentário de Prager em 1924 com a combinação de intertítulos e títulos na tela.....	14
Figura 6 - Duas fotos do documentário anti-guerra Shanghai (1938), feito por Fumio Kamei, mostrando legendas em japonês exibidas na lateral da tela.....	16
Figura 7 - Estenotipista criando Closed Caption em teclado personalizado para a função.....	25
Figura 8 - Espectro de potência de fragmentos das vogais quando falada.....	28
Figura 9 - Vídeo de demonstração de como utilizar a ferramenta Happy Scribe.....	39
Figura 10 - Vídeo de demonstração de como utilizar a ferramenta Rev.....	40
Figura 11 - Vídeo de demonstração de como utilizar a ferramenta Amberscript.....	41
Figura 12 - Vídeo de demonstração de como utilizar a ferramenta Transcribe.....	42
Figura 13 - Vídeo de demonstração de como utilizar a ferramenta Trint.....	43

Figura 14 - Vídeo de demonstração de como utilizar a ferramenta Temi.....	44
Figura 15 - Vídeo de demonstração de como utilizar a ferramenta Descript.....	45
Figura 16 - Vídeo de demonstração de como utilizar a ferramenta Sonix.....	45
Figura 17 - Vídeo de demonstração de como utilizar a ferramenta Audext.....	46
Figura 18 - Linha de código de importação da biblioteca de transcrição.....	52
Figura 19 - Códigos do módulo de apoio que fazem o background da transcrição.....	52
Figura 20 - Códigos que efetuam de fato a transcrição do áudio em texto.....	53
Figura 21 - Protótipo ISA em funcionamento, teste de transcrição via Microfone.....	53
Figura 22 - Protótipo ISA em funcionamento, teste de transcrição a partir de vídeo do Youtube via alto-falante.....	53

LISTA DE QUADROS

Quadro 1 - Cronologia das últimas etapas executadas.....	05
--	----

LISTA DE TABELAS

Tabela 1 - Tabela comparativa entre o ISA e os outros softwares transcritores de texto.....	36
---	----

LISTA DE ABREVIÇÕES E SIGLAS

ADC - Analogic Digital Converter - Conversor Analógico-digital
ANN - Artificial Neural Network
API - Application Programming Interface
ATV - Audiovisual Translate Version - Tradução Audiovisual
BLS - Bureau of Labor Statistics
CART - Communication Access in Real Time
CEO - Chief Executive Office
CES - Consumer Electronics Show
CTC - Connectionist Temporal Classification
DCNC - Digital Cinema Naming Convention
EDCF - European Digital Camera Forum
EN - English - Idioma Inglês
EUA - Estados Unidos da América
FR - French - Idioma Francês
HMM - Hidden Markov Model
IA - Inteligência Artificial
ISA - Interface de Suporte à Acessibilidade
LIBRAS - Língua Brasileira de Sinais
LSE - Legendagem para Surdos e Ensurdidos
LSTM - Long Short-Term Memory
LVT - Laser Vidéo Titres
MIT - Massachusetts Institute of Technology
ML - Machine Learning
RA - Realidade Aumentada
RNA - Redes Neurais Artificiais
SEO - Search Engine Optimization
STT - Speech-to-Text
VAD - Voice Activity Detector

SUMÁRIO

1 RELATÓRIO TÉCNICO-CIENTÍFICO	17
1.1 Dados Gerais do Projeto	17
1.2 Objetivo Geral	18
1.3 Objetivos Específicos	18
2 JUSTIFICATIVA	19
3 ÚLTIMAS ETAPAS EXECUTADAS	20
4 DIFICULDADES ENCONTRADAS	20
5 REFERENCIAL TÉCNICO-TEÓRICO	21
5.1 CLOSED CAPTION; O que é o Closed Caption	21
5.2 CLOSED CAPTION NO BRASIL	34
5.3 PYTHON — LINGUAGEM DE PROGRAMAÇÃO	39
5.3.1 História do Python	39
5.3.2 Python na Transcrição de Áudio em Texto	40
5.3.3 Bibliotecas de Transcrição de Áudio em Texto que Utilizam Machine Learning	42
5.3.4 História do Machine Learning	43
5.3.5 Machine Learning na Acessibilidade	49
5.4 SOFTWARES TRANSCRITORES DE ÁUDIO EM TEXTO	51
5.4.1 Happy Scribe	52
5.4.2 Rev	52
5.4.3 Amberscript	53
5.4.4 Transcribe	54
5.4.5 Trint	55
5.4.6 Temi	56
5.4.7 Descript	57
5.4.8 Sonix	58
5.4.9 Audext	58
5.4.10 Comparação entre softwares de transcrição de áudio em texto	59
5.5 ISA — INTERFACE DE SUPORTE À ACESSIBILIDADE	61
5.5.1 ISA versus LSE/Closed Caption	62
5.5.2 Benefícios Gerais da ISA	63
6 CÓDIGO FONTE	63
7 CONCLUSÃO	65
8 REFERÊNCIAS	66

1 RELATÓRIO TÉCNICO-CIENTÍFICO

1.1 Dados Gerais do Projeto

Nome do Produto:

ISA (Interface de Suporte à Acessibilidade), Protótipo de Software de legendagem compatível com computadores pessoais e produtos audiovisuais.

Período de Execução:

8(oito) meses.

Valor total do projeto:

O projeto não recebeu investimento financeiro direto. Porém, teve o apoio humano e de infraestrutura oferecidos pela Unesp campus de Bauru, além de recursos de hardware (notebook) anteriormente adquiridos pelo pesquisador.

Bolsas — Financiamentos — Convênios e Parcerias:

Não houve.

Instituições participantes:

O projeto foi desenvolvido nas dependências da UNESP – Universidade Estadual Paulista “Júlio de Mesquita Filho” campus de Bauru, durante o curso de mestrado do PPGMiT – Programa de Pós-Graduação em Mídia e Tecnologia vinculado a FAAC – Faculdade de Arquitetura, Artes, Comunicação e Design e na residência do pesquisador em diversos momentos de atividade remota. Para que fosse possível a implementação do protótipo, foi necessária a utilização de recursos de hardware, software e conexão com a internet, que foi disponibilizada nesta parceria com o Programa de mestrado

Caracterização da Pesquisa:

Pesquisa Técnico Científica de Base Tecnológica, reproduzível e aplicável.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

Caracterização da pesquisa com uma breve justificativa para o enquadramento:

O Brasil possui quase 11 milhões de deficientes auditivos (Raio-X da Surdez no Brasil, 2019), que representam cerca de 5% da população brasileira, sendo que mais de dois milhões deles têm deficiência em seu nível severo, o que significa que não ouvem absolutamente nada, ou tem até 20% de audição.

Atualmente as tecnologias assistivas para TV no Brasil, direcionada para o surdo está longe de serem ideais. Existem, mas são bastante ineficazes em relação ao atraso contido na transcrição, que deixa uma diferença de aproximadamente cinco a dez segundos entre a legenda descritiva e a imagem exibida, prejudicando o entendimento do conteúdo por parte dos surdos e ensurdecidos, além de serem caras demais para organizações que possuem pouca ou nenhuma renda proveniente dos seus produtos audiovisuais, como por exemplo, TVs universitárias ou comunitárias, produtoras independentes e de conteúdo individuais para a internet.

A mais conhecida delas é a "Legenda Descritiva para TV" denominada de "*Closed Caption*", que apresenta a descrição da cena por "Linguagem Fácil", porém tem um alto custo de implementação e manutenção. Por exemplo, uma implementação dessa ferramenta em sua versão simples, beira a casa dos noventa mil reais de custo, tornando-se inviável para as pequenas e médias redes de comunicação, que por esse motivo deixam de atingir essa parcela da população.

Através do Decreto n.º 5.296, de 2 de dezembro de 2004, ficou estipulado que equipamentos, sistemas de comunicação e informação aos deficientes físicos devem ser utilizados para garantir acessibilidade, tornando assim obrigatória a utilização do *Closed Caption* em programação de rede aberta e impossibilitando o uso dela por instituições e redes de comunicação universitárias ou sem o devido orçamento (BRASIL, 2004). Sendo assim, esta pesquisa buscou estudar meios para melhorar e viabilizar o uso da ferramenta por essas organizações, principalmente através das mais altas tecnologias disponíveis — Modelos de Inteligência Artificial, técnicas de aprendizado de

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

máquina, armazenamento e requisição de dados em nuvem, visando o melhor custo-benefício para a implementação dela.

Pesquisador:

Felipe Rodolfo Domingues da Silva

Orientador:

Prof. Dr. Francisco Machado Filho

Equipamentos adquiridos:

Para o desenvolvimento do projeto foi necessária a utilização de um notebook de modelo Macbook Air M1 que já era de propriedade do pesquisador antes do ingresso na pós-graduação, além do acesso à internet.

Link de disseminação em repositório de código aberto:

<https://github.com/domfelipe/ISA-Interface/tree/main/Download>

1.2 Objetivo geral

O objetivo desta pesquisa é a partir das tecnologias como bibliotecas de transcrição de áudio suportadas pela linguagem de programação Python e módulos selecionados de inteligência artificial, desenvolver um protótipo de um *software* de transcrição de áudio em texto em tempo real, assemelhado ao *Closed Caption*, melhorando seu tempo de resposta, a assertividade da fala junto ao receptor e tornando acessível do ponto de vista financeiro, para organizações que possuam pouco ou nenhum recurso para disponibilizar seus produtos audiovisuais.

1.3 Objetivos específicos

Os objetivos específicos deste projeto são:

1. Analisar o desempenho e a usabilidade do modelo utilizado em 2021/2022 da ferramenta *Closed Caption* para uso pessoal de surdos e ensurdecidos.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

2. Criar um protótipo de código aberto para usuários ensurdecidos priorizando o custo-benefício.
3. Fazer com que esse protótipo seja a interface, que ligará o público surdo à acessibilidade em produtos audiovisuais de livre escolha, desse modo, o usuário poderá consumir o conteúdo audiovisual de seu desejo e ter a assistência acessível a todo instante.
4. Desenvolver um processo ou passo a passo para o uso mais eficiente de tecnologias digitais de transcrição de áudio.
5. Testar e aplicar o protótipo da matriz indicada numa organização (SEM FINS LUCRATIVOS) de nenhuma renda proveniente dos seus produtos audiovisuais, como por exemplo TVs universitárias, TVs comunitárias, produtoras independentes e produtores de conteúdo individuais para a internet.

2 JUSTIFICATIVA

Segundo identificado em pesquisas realizadas pelos órgãos de defesa da inclusão da pessoa surda, como o Instituto Locomotiva (Raio-X da Surdez no Brasil, 2019) — a mais recente em 2019, a parcela da população acometida pela surdez cresce em um ritmo contínuo. No Brasil temos cerca de 10,7 milhões de pessoas que apresentam algum nível de deficiência auditiva, sendo que 21,8% desses tem severa deficiência, o que significa que eles têm apenas 20% ou menos de audição. E esse índice deverá aumentar até 2050, com o crescimento populacional e a falta de cuidado da população com a própria audição.

De igual forma é detectado que as tecnologias assistivas direcionadas a essas pessoas não acompanham os avanços tecnológicos junto a esse crescimento, pelo fato de que, as ferramentas de desenvolvimento que se adaptam à essas tecnologias surgiram a pouco tempo, trazendo a possibilidade da concepção de melhorias e aperfeiçoamentos destas atualmente disponíveis e criação de outras novas apenas recentemente.

Entretanto, outro fator limitante da utilização das tecnologias assistivas para o audiovisual é o alto custo de implantação e manutenção delas, visando o

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

público geral, produtores de conteúdo e pequenas redes de comunicação. Por esse motivo esta pesquisa busca estudar meios para melhorar e viabilizar o uso da ferramenta - *Closed Caption* - por essas organizações, a partir da análise de desempenho e do custo-benefício, e da criação de um protótipo de baixo custo, possibilitando, a fácil implementação desta, auxiliando na disseminação e fomentando do uso de ferramentas assistivas.

3 ÚLTIMAS ETAPAS EXECUTADAS

Com o desenvolvimento em fase avançada do protótipo, iniciou-se a fase de aprimoramento dele, visando a compatibilidade para com qualquer categoria de computador. Essa fase teve início em outubro e se findou em novembro com o período em que foram feitos os testes finais deste em diferentes máquinas e configurações. Estes testes ocorreram em paralelo com a revisão atual do projeto e da análise dos resultados, além da adição do código-fonte explicado e exemplificado.

Após o término de inserção de todos dados, resultados e informações, foi executada a revisão geral do relatório técnico visando a finalização da adequação de textos às normas ABNT, além do intuito de encontrar, porventura, algum erro no projeto.

Finalizando o processo de desenvolvimento do projeto, teve início a sua redação final em janeiro, seguido pela sua entrega para a defesa.

ETAPAS	10/2021	11/2021	12/2021	01/2022	02/2022	03/2022
Revisão do Projeto						
Adição dos Resultados						
Revisão Geral						
Redação Final						
Entrega						

Quadro 1 - Cronologia das últimas etapas executadas.

4 DIFICULDADES ENCONTRADAS

- Por conta do cenário de pandemia, não conseguimos efetuar os testes de implementação em uma estrutura de TV oficial.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

- Por mais que tenhamos tecnologias avançadíssimas como a inteligência artificial em seu campo de “Deep Learning” que eleva o nível de aprendizado e reconhecimento de um robô construído e treinado, não temos um módulo de IA totalmente grátis, para encaixar perfeitamente ao projeto.
- A literatura sobre o assunto: “Legendagem para surdos e ensurdecidos” em português e de autores brasileiros ainda é escassa, porém, está crescendo aliada a movimento de igualdade e acessibilidade que vem ganhando destaque nacionalmente.
- Nem tudo foi ruim nesse período pandêmico, para mim, particularmente, foi propício, pois pude ficar home-office. Com isso, concluí todas as matérias e créditos, o que seria quase inviável se estivéssemos em período de aulas e trabalho presencial.

5 REFERENCIAL TÉCNICO-TEÓRICO

5.1 CLOSED CAPTION

O que é o Closed Caption

As legendas ocultas são tão comuns que muitas vezes as consideramos garantidas — mas nem sempre foi esse o caso. Como todo produto tecnológico, elas evoluíram até o seu estado atual e se tornaram essenciais para a acessibilidade e para o público deficiente auditivo. Entretanto, temos de dissipar alguns equívocos comuns em relação aos tipos de legendas.

Nem mesmo a Netflix diferencia completamente os tipos de legendas, por isso, não é surpresa que a maioria das pessoas pense desta maneira. Por um lado, tanto as legendas ocultas quanto às padrão (aquelas que vemos em produtos audiovisuais de formato “Legendado”) compartilham uma história: ambas são versões da Tradução Audiovisual (ATV) na modalidade legendagem,

O modo de tradução audiovisual mais amplamente estudado da legendagem é a prática linguística que mostra o texto escrito em uma tela que transmite "uma versão do idioma de destino do discurso de origem". Consistindo em muitos subtipos, o mais

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

comumente utilizado é a legendagem interlinguística, que geralmente é exibida em legendas abertas. Em locais onde vários idiomas são falados, as legendas bilíngues são usadas para mostrar duas versões de idioma diferentes do texto de origem ao mesmo tempo (PÉREZ- 2014, p.15).

Mas existem algumas diferenças importantes entre eles. As legendas padrão hoje são traduções para públicos que não falam o idioma original do vídeo. Eles incluem apenas diálogo e não atendem aos requisitos de acessibilidade para deficiência auditiva. Eles também são geralmente criados antes do lançamento do produto audiovisual.

As legendas ocultas (*Closed Caption*), também conhecidas como legendagem para surdos e ensurdecidos (LSE) existem para tornar o áudio de um produto audiovisual acessível ao público com necessidades auditivas. Assim, eles incluem descrições de diálogo e ruído de fundo. Eles podem ser abertos (permanentes) ou fechados (o espectador pode ativá-los ou desativá-los).

Hoje em dia, estamos todos familiarizados com os benefícios das legendas ocultas para surdos e deficientes auditivos. Mas muitas pessoas não percebem que eles oferecem benefícios significativos para uma infinidade de outras razões. Por um lado, eles ajudam a tornar seus produtos audiovisuais mais visíveis nas ferramentas de buscas da internet. Isso porque as legendas ocultas permitem que os mecanismos de pesquisa indexarem para a Search Engine Optimization(SEO), efeito esse que só pode ser realizado se houver algum tipo de texto para rastrear no filme ou vídeo.

As legendas também melhoram o tempo médio de exibição de seus vídeos ao envolver seu público. E não apenas isso, um educador ou empregador que produz vídeos, também se beneficia, pois, além de melhorar a acessibilidade, as legendas melhoram os resultados da aprendizagem. E algumas outras situações que exigem legendas.

História do *Closed Caption*

Até hoje há divergência no ecossistema audiovisual sobre quando os termos 'intertítulos' e 'legendas' surgiram pela primeira vez, embora a origem do primeiro possa ser mais recente do que muitos supõem e para isso

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

precisamos efetuar uma verdadeira viagem cronológica cinematográfica para explicar o fato.

O professor André Gaudreault afirma que o dicionário Roberts remonta o 'intertítulo' francês até 1955, enquanto os dicionários Oxford afirmam que o 'intertítulo' inglês tem origem na década de 1930. (Flycz Blog, 2014) Como que para turvar ainda mais as águas, o termo hifenizado 'sub-títulos' data de 1826, em seu contexto literário, embora haja evidências que 'legendas' sempre foi o termo padrão para o que são geralmente chamados de intertítulos (SLIDE,2001). Então, embora pareça que essa explicação pode ser confusa, pois — 'intertítulos' é um retrônimo — neologismo gerado pela modificação do nome original de um objeto/conceito para diferenciá-lo da versão mais recente dele — (assim como 'filmes mudos'). Todavia, por uma questão de continuidade, para evitar confusão, ele será usado quando nos referirmos aos cartões de título inseridos da era do 'filme mudo'.

Auguste e Louis Lumière, Georges Méliès; Thomas Edison — são alguns dos nomes famosos que geralmente vêm à mente quando se pensa no início da produção cinematográfica, enquanto aqueles que criaram e desenvolveram a legendagem ainda permanecem na maioria desconhecidos fora do mundo do cinema mudo profissional. Por exemplo: uma variação do Edison Kinetoscope (2021) foi fabricada pelo inovador e visionário cineasta britânico Robert W. Paul (BARNES, 1976) que, em abril de 1896, trazendo pela primeira vez a possibilidade de adicionar intertítulos (letras ou frases) em uma produção cinematográfica, Paul incluiu este engenhoso novo dispositivo em sua 'Câmera Cinematográfica No.1' e logo provou seu valor em relação à titulação na tela.

No entanto, Robert também inventou intertítulos? Essa questão ainda é motivo de debate entre os estudiosos do cinema primitivo, pois o cartunista e cineasta americano nascido na Grã-Bretanha, James Blackton pode ter utilizado intertítulos em 1897, um ano inteiro antes de Robert estrear o primeiro seu filme oficialmente registrado a usar intertítulos (The British Film Catalogue, [s.d.]). Infelizmente, há escassas evidências físicas sobreviventes desta época para provar isso. No entanto, ainda em 1898, Robert exibiu “*Our*

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

New General Servant”, drama leve com legendas, com poucos minutos de duração e apenas quatro cenas, em que uma mulher descobre o marido beijando a empregada e a dispensa prontamente. Ele é considerado um dos primeiros produtos cinematográficos com intertítulos.

Ao amanhecer do século 20, os experimentos com intertítulos continuaram. No ano de 1900 outro cineasta britânico, Cecil M. Hepworth (HEPWORTH, 1951) apresentou seu curta-metragem cômico *“How it Feels to Be Run Over”*, no qual o operador de câmera parece realmente ser atropelado por uma daquelas carruagens sem cavalos modernos; os ainda impopulares automóveis, e então o filme termina com uma curiosa sequência de títulos escrita à mão que diz simplesmente “Oh! Mamãe ficará feliz!”, o que sugere que esse seja o primeiro intertítulo oralizado utilizado no cinema (NEILSON; MANVELL; WOLLENBERG, 1948), conforme a figura 1.

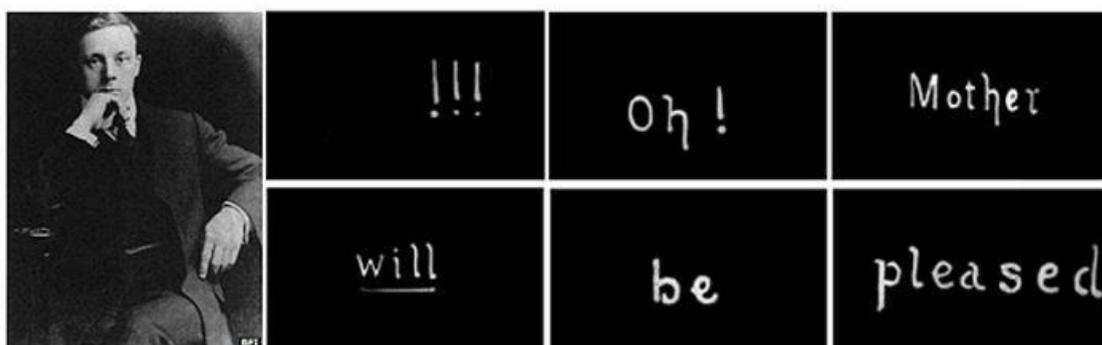


Figura 1 - Cecil M. Hepworth e os primeiros intertítulos oralizados do cinema mundial. Fonte: How It Feels to Be Run Over, Mubi.com [s.d.]

Um ano depois, o mecanismo de Robert Paul foi colocado em uso efetivo em sua produção de 1901 de *“Scrooge, or Marley’s Ghost”*, integrando a palavra ‘Scrooge’ na cena de abertura e uma técnica semelhante foi empregada por Hepworth para seu filme *“Alice no País das Maravilhas”*. No mesmo ano, nos Estados Unidos, o Kinetoscope foi utilizado em *“Uncle Tom’s Cabin”* de Edwin S. Porter (frequentemente erroneamente identificado como o primeiro filme intertítulo do mundo) e em 1907, *“Porter’s College Chums”* apresentou um uso verdadeiramente imaginativo do que, no mundo moderno, poderia ser chamado de ‘mensagens de texto verbais’, superando Hepworth de maneira inovadora, durante uma conversa telefônica na qual uma jovem repreende

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

seu noivo após vê-lo com outra mulher, como podemos observar na figura 2.



Figura 2 - Os intertítulos utilizados por Robert Paul em 'Scrooge, or, Marley's Ghost' (1901) e a utilização de Hepworth 'Alice no País das Maravilhas' (1903). Fonte: The First Film to Use Intertitles Was Also the First Dickens Adaptation, ETTLEMAN, 2017; "Alice in Wonderland (Film, Fantasy): Reviews, Ratings, Cast and Crew - Rate Your Music," [s.d.]

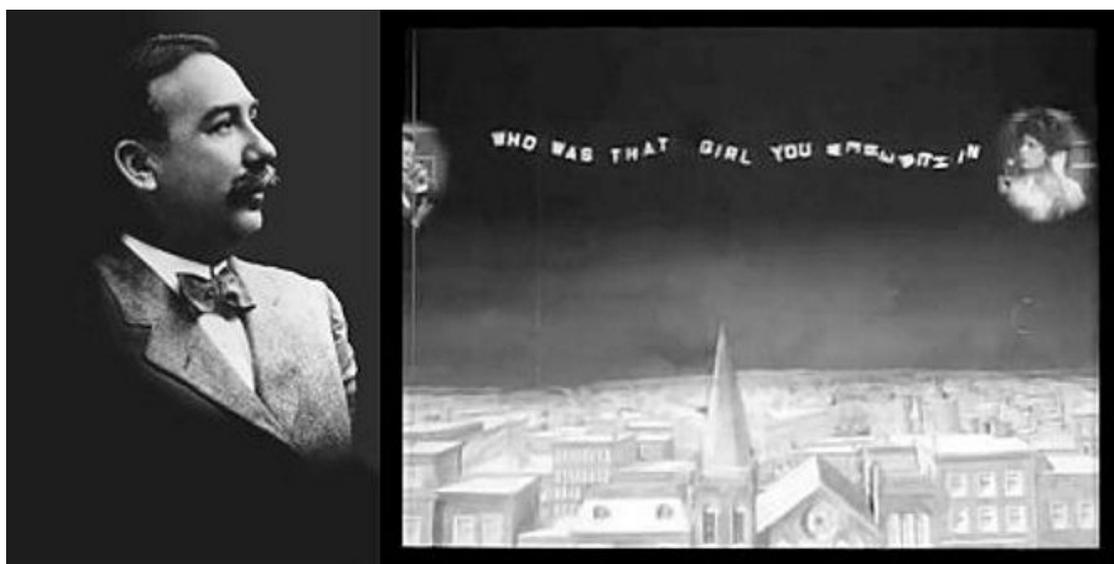


Figura 3 - Edwin S. Porter e seus 'textos verbais' em 'College Chums' (1907). Fonte: The first movie with verbal texting - 'College Chums' 1907 by Edwin S Porter Thomas Edison," [s.d.]

Sem alarde ou pompa, os intertítulos aumentaram gradualmente a sua presença na década seguinte, sendo frequente em uma época considerada de inovação quase constante; o que levou a que a titulação geral se tornasse cada vez mais elaborada em seu design e uso de texto. Os intertítulos expositivos e de diálogo tornara-se o padrão — o primeiro explicando os pontos da trama, o segundo ilustrando as falas dos atores. As sequências de títulos de abertura começaram a listar os nomes das empresas de produção e creditaram o diretor e as estrelas do filme; e com acompanhamento musical — piano, órgão, quarteto de cordas, até mesmo uma orquestra completa às

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

vezes, 'ir ao cinema' deu um passo mais próximo da experiência contemporânea.

Com o surgimento de novas produtoras nos Estados Unidos e na Europa, surgiu a inevitável questão da tradução para os mercados de exportação. Os intertítulos não apresentavam nenhum obstáculo, pois podiam ser facilmente removidos, traduzidos, filmados e inseridos novamente, enquanto às vezes o problema de tradução era resolvido com atores re-dublado ou 'dublado ao vivo' o diálogo por trás da tela. Na França e no Japão, especialmente filmes americanos, seriam exibidos com um comentarista interpretativo acompanhante no palco – o “*Bonimenteur*” ([s.d.]) na França e “*The Benshi*” (“THE BENSHI TRADITION: CINEMA=PERFORMANCE,” [s.d.]) no Japão. Este interlocutor explicaria toda a história e traduziria os intertítulos, além de interpretar a ação na tela e no Japão em particular, dublaria criativamente o diálogo silencioso dos atores.

No entanto, desde 1909 foram feitas tentativas para complementar os filmes mudos com o que um público moderno reconheceria como legendas. Os primeiros experimentos conhecidos (FONG, 2009) incluíam o uso de um sciopticon (semelhante a uma lanterna) para projetar texto na tela; um método que dependia de o projetorista exibir slides de vidro precisamente no tempo com o diálogo silencioso dos atores. Uma abordagem semelhante imprimiu o texto em uma tira de filme contígua, apelidada de 'títulos de linha' na Dinamarca (HENRIK, 2005), embora mais uma vez isso dependesse de sincronização absoluta, mas como se para agravar ainda mais o problema, o texto projetado (slide ou tira de filme) era muitas vezes difícil de ler.

Uma ideia para iluminar o diálogo na tela, que não se encaixa na categoria de legendagem ou legendagem, foi concebida por Abraham S. Schomer (1920) que tentando romper com as limitações da legendagem, ele escreveu, dirigiu e co-produziu “*The Chamber Mystery*” (“Silent Era: Progressive Silent Film List,” [s.d.]), lançado em 1920, que combinava intertítulos com balões de palavras no quadro adicionados na pós-produção. Infelizmente para Schomer, esses complementos de quadrinhos não conseguiram capturar a imaginação do público espectador e seu filme é amplamente ignorado na história do

cinema; a ironia é que o texto gerado por computador moderno pode ser programado para aparecer em qualquer lugar na tela e até rastrear os movimentos de um personagem, então talvez Schomer (1920) não fosse tão excêntrico assim, mas às vezes um visionário incompreendido.



Figura 4 - Dois frames de “The Chamber Mystery” (1920) utilizando a técnica de balões de palavras. Fonte: “Silent Era : Progressive Silent Film List,” [s.d.]

Na época em que Schomer estava experimentando com seus balões de palavras Herman Weinberg, estava reorganizando partituras sinfônicas de filmes mudos alemães. Ele também começou a traduzir as primeiras imagens faladas da Alemanha, o que o levou a uma nova carreira na qual ele legendou mais de 300 filmes e escreveu extensivamente sobre cinema. Em entrevista a Weinberg, sobre seus dias como legendador, ele comentou:

No início, tentamos a técnica usada para fotos silenciosas. A cada poucos minutos haveria um título em tela cheia. Mas isso não funcionava porque sempre havia algumas pessoas que entendiam alemão e riam das piadas faladas no filme. Todos os outros ficaram irritados porque achavam que estava faltando alguma coisa. (A WEE CORNER OF SHOWBIZ, 1960)

Mais tarde, depois de adquirir uma máquina que permitia ao editor ver o filme durante a edição, Weinberg (1982, p. 108) testou cautelosamente seu potencial de exibição de texto dentro das imagens em movimento, na parte inferior da tela, observando que o público “não baixou as cabeças, eles apenas baixaram os olhos” (WEINBERG, 1982).

Uma história divertida é contada sobre o humor seco de Weinberg e sua abordagem engenhosa para contornar a censura, anos mais tarde em

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

regimes ditatoriais na Alemanha. Deliberando sobre como traduzir um comentário particularmente ofensivo em um filme francês e sabendo que a tradução exata nunca seria aprovada pelos censores americanos, Weinberg, como Schomer antes dele, encontrou inspiração na abordagem dos quadrinhos para xingamentos e sua solução foi utilizar - “XZ%!X” - símbolos e letras aleatórios. (FREEDMAN, 1983)

Voltando alguns anos para 1924, o produtor de cinema alemão Wilhelm Prager lançou um documentário esportivo “*Wege zu Kraft und Schönheit*” (*Ways to Strength and Beauty*), que apresentava uma combinação de intertítulos e títulos na tela para adicionar comentários e nomear os atletas concorrentes. (“*Ways to Strength and Beauty*,” 2022) Embora não fosse uma ideia nova na época, isso, no entanto, indicava um movimento predominante da legendagem, como vemos abaixo na figura 5.



Figura 5 - Documentário de Prager em 1924 com a combinação de intertítulos e títulos na tela. Fonte: Youtube (“N. Kaufmann & W. Prager: *Wege zu Kraft und Schönheit* (1925),” [s.d.])

Com as traduções de Weinberg, os títulos na tela de Prager e o lançamento

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

em 1927 do filme semi-falado da Warner Brothers, *“The Jazz Singer”* (CROSLAND et al., 1928) – a escrita estava finalmente na caixa para intertítulos. À medida que os filmes falados de Hollywood começaram a dominar o cenário cinematográfico mundial, esse novo fenômeno deu um impulso muito necessário ao aprimoramento da legendagem para os mercados estrangeiros. Com base nas técnicas usadas nos documentários de Prager e nos filmes americanos, como o veículo estelar de Clara Bow de 1925, *“The Plastic Age”* (RUGGLES et al., 1925), foi desenvolvido um processo óptico, no qual os quadros legendados únicos eram mantidos em posição enquanto o negativo do filme e a tira de impressão positiva eram alimentadas para a frente e exposta ao mesmo tempo. O resultado foi uma melhoria na legibilidade das legendas e em 1929, *“The Jazz Singer”* (CROSLAND et al., 1928), dessa vez com legendas em francês, estreou em Paris, seguido por um lançamento legendado em dinamarquês de *“The Singing Fool”* (BACON et al., 1929), também em 1929. A resposta do público foi, no entanto, pouco entusiasmada no início, sendo as legendas consideradas 'irritantes' por alguns (NORNES, 2007).

Enquanto isso, na terra do sol nascente, dois mestres da legendagem japonesa, Shimizu Shunji e Tamaru Yukihiro, faziam experiências com legendas projetadas na lateral da tela. Embora Shimizu seja geralmente creditado por inaugurar a legendagem adequada no Japão. Foi Yukihiro, quem criou as legendas para o filme de Josef Von Sternberg de 1930, *“Morocco”* (STERNBERG et al., 1930), lançado no Japão em 1931, o primeiro filme de Hollywood legendado para um público japonês doméstico. Uma opinião amplamente difundida entre os legendadores japoneses nativos da época, expressa aqui pelo contemporâneo de Shunji, Okaeda Shinji (NORNES, 2007, p.21), era que “Quanto menos palavras um filme tiver, melhor(sic)”, essa opinião também se aplicava a músicas estrangeiras. Yukihiro apresentou Shimizu à Paramount Pictures, que o contratou para preparar cinejornais americanos legendados para distribuição em todo o Japão durante as décadas de 1930 e 1940, até que os filmes americanos foram finalmente proibidos no Japão com a eclosão da Segunda Guerra

Mundial.



Figura 6 - Duas fotos do documentário anti-guerra Shanghai (1938), feito por Fumio Kamei, mostrando legendas em japonês exibidas na lateral da tela. Fonte: Youtube ("A Record of Shanghai (1938)," [s.d.]

Enquanto Hollywood estendeu seu domínio sobre a indústria cinematográfica dos anos 30 aos anos 50 com a *'Golden Age'* do Cinema, na Europa, durante esse mesmo período, os avanços nas técnicas de legendagem foram desenvolvidos na Noruega, Suécia, Hungria e França. Tudo começou em 1930, quando um inventor norueguês chamado Leif Eriksen patenteou um método mecânico para estampar texto diretamente na tira de filme. Citando o artigo do escritor sueco Jan Ivarsson, 'A Short Technical History of Subtitles In Europe',

[...]primeiro umedecendo a camada de emulsão para amolecê-la. Os títulos foram compostos, impressos em papel e fotografados para produzir chapas tipográficas muito pequenas para cada legenda (a altura de cada letra sendo apenas cerca de 0,8 mm), que seria então pressionada na camada de emulsão para criar a legenda desejada. (IVARSSON, 2004)

Então, em 1932, um processo químico foi patenteado simultaneamente pelos inventores húngaros e noruegueses, R. Hruška (em Budapeste) e Oscar I. Ertnæs (em Oslo), usando chapas de impressão aquecidas pressionadas contra o lado da emulsão de um revestimento de cera ou parafina, cópia finalizada do filme. O revestimento derreteu e expôs a emulsão por baixo. O

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

filme foi então lavado com

água sanitária para dissolver a emulsão exposta, produzindo letras brancas legíveis quando projetadas em uma tela, embora as bordas fossem levemente irregulares devido a variações no processo (IVARSSON, 2004)

Três anos depois, um processo térmico foi patenteado por outro húngaro, O. Turchányi, “pelo qual as placas eram aquecidas a uma temperatura suficientemente alta para derreter a emulsão no filme sem a necessidade de um banho de amolecimento”, mas havia um problema com este processo, como com a abordagem anterior de Eriksen, ambos eram difíceis de controlar e as legendas resultantes nem sempre eram legíveis. Durante o final da década de 1940, na Grã-Bretanha, a Rank Organization (“The Rank Organisation,” 2022) experimentou brevemente com legendas gravadas em placas de vidro que foram exibidas, através de um projetor, no canto inferior esquerdo da tela do cinema que embora fosse uma versão melhorada do processo de Turchányi, e os processos químicos anteriores de Hruška e Ertnæs, continuaram a serem usados na Europa e também permaneceu sendo a legendagem dominante em partes da Ásia e América do Sul até o final dos anos 1990. (IVARSSON, 2004)

Em seu site, a legendadora francesa Maï Boiron descreve suas funções durante os dias da legendagem química no início dos anos 1990. “Usando um microscópio, eu revisava cada pequena placa de zinco usada para gravar as legendas no filme, quadro a quadro” (“BIO | Maï Boiron | Sous-titres et doublage,” [s.d.]). Uma tarefa que exigia um olhar muito treinado e preciso. Mais tarde, Boiron se juntaria à empresa Titra-Film em Paris, que, junto com o empresário da Laser Vidéo Titres (LVT), Denis Auboyer, vinha desenvolvendo legendagem gravada a laser desde meados da década de 1980. No entanto, houve problemas iniciais durante os testes da nova tecnologia, com o laser queimando tanto a emulsão quanto o filme e, ainda assim, com o ajuste, a 'queima' acabou sendo uma vantagem. Citando Auboyer,

É como se você tivesse um cigarro e um pedaço de papel branco. Você faz um buraco no papel com o cigarro e ao redor dele há um círculo preto. Isso é o que temos em torno das legendas, para que você possa vê-las mesmo quando o fundo é branco. (AUBOYER,

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

[s.d.]

O cartão de visita de Auboyer foi a legendagem do filme “*Bird*” de Clint Eastwood para o festival de Cannes de 1988 por LVT, afirmando que “Embora este não tenha sido nosso primeiro filme, nós o consideramos um dos mais importantes para nossa reputação” (AUBOYER; CATELLANI, [s.d.]) e, portanto, os antigos métodos de composição manual, químicos banhos e placas de aquecimento tornaram-se finalmente redundantes no mercado europeu. Hollywood adotou o mesmo processo a laser e, na década de 1990, a titulação a laser tornou-se o padrão da indústria, com a legendagem para o público doméstico americano predominantemente feita por empresas em Los Angeles e Londres.

O desenvolvimento de novas tecnologias é incessante, como a história até agora mostrou e enquanto Maï Boiron revisava aquelas minúsculas chapas de zinco, uma revolução estava em andamento na América. O ano de 1992 viu a primeira demonstração pública de cinema digital; e em 2001 foi criado o Fórum Europeu de Cinema Digital (EDCF) (“EDCF | Guide to Digital Cinema Mastering,” [s.d.]) para discutir a progressão da tecnologia digital e, em particular, a legendagem de alta qualidade.

As legendas são geralmente o último pensamento na mente de qualquer pessoa ao planejar um pacote de mídia para distribuição em massa... A necessidade de contar histórias precisas, inteligentes e convincentes é de suma importância... [mas]... a característica desanimadora da legendagem é o custo. Há um prêmio de serviço externo que não pode ser evitado; uma experiência que não pode ser replicada por sites como tradutores online. (EDCF | Guide to Digital Cinema Mastering, [s.d.] p.18)

Dois anos depois, em 3 de novembro de 2003, a Texas Instruments lançou seu 'Subtitle Specification (XML Format) for DLP Cinema Projection Technology' e nos dez anos seguintes, vinte patentes de formato de legenda digital foram solicitadas na América, Europa e Japão (Instituto Europeu de Patentes, Pedido de Patente EP 2 938 087 A1. p.7-8). Em 2010, a Digital Cinema Naming Convention (DCNC) (“Illustrated Guide,” [s.d.]) estava desenvolvendo codificação para pacotes de filmes de cinema digital para identificar, entre outros aspectos, o elemento de legendagem. Um exemplo é Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

o seguinte: “*TÍTULO DO FILME-GL_FTR_F_FR -EN _FR-GB_51-FR_2K_ST_20070115_FAC-i3D_OV*”.

O que parece uma confusão sem sentido de letras e números para uma pessoa comum pode ser decifrada de relance pelo projetorista receptor. Uma explicação completa da codificação DCNC pode ser encontrada na documentação disponível ao público (Illustrated Guide [s.d.]), mas apenas os caracteres 'FR-EN' são diretamente relevantes, pois indica que o filme de amostra está em francês (FR) com legendas em inglês (EN).

Hoje, existem dois métodos para incluir legendas em um filme digital. O primeiro tem legendas 'abertas', codificadas no formato digital e, não podem ser removidas ou desligadas. O segundo tem legendas 'fechadas', incluídas como um arquivo separado, portanto, devem ser selecionadas através do software do computador e, em seguida, o projetor digital sincroniza automaticamente as legendas com o filme.

Com o recente interesse de desenvolvimento em Realidade Aumentada (RA), talvez não seja surpreendente que, mais cedo ou mais tarde, alguém, em algum lugar, possa fazer de um simples par de óculos uma ferramenta de exibição e tradução de legendagem, assim como em 2009 na Sony Consumer Electronics Show, o ator Tom Hanks demonstrou o protótipo da Sony para óculos de legenda. Em 2012, os óculos estavam passando por testes generalizados em cinemas na América e na Europa, enquanto no Japão a Seiko Epson Corporation já tinha desenvolvido um produto similar em 2011 (“User’s Guide,” [s.d.]). Embora o objetivo final deste último seja o mercado de RA, seus primeiros testes de legendagem foram realizados em parceria com a Japanese New National Theatre Foundation em 2014. Ambas as invenções usam Wi-Fi para receber um sinal do servidor do teatro e exibir as legendas na superfície das lentes.

Enquanto isso, na Grã-Bretanha, durante 2013 e 2014, o ex-professor que virou inventor, Jack Ezra, estava desenvolvendo o que ele descreveu como um 'Sistema de Legendas de Cinema Off-Screen' (“Invisible Subtitles,” [s.d.]) para surdos e deficientes auditivos. Em comparação com os produtos da Sony e da Epson (caros, facilmente danificados em caso de queda e supostamente

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

desconfortáveis quando usados por longos períodos), o sistema da Ezra é único e consideravelmente mais barato.

“Nosso sistema coloca uma tela especial 'discreta' na parte inferior da tela do filme, que parece em branco. Assim, não distrai o público em geral[] Para quem quiser ver as legendas, basta colocar um leve par de óculos, que permite ao espectador ler as legendas em nosso display, que mostra legendas grandes e claras em sincronia ao filme” (“Off-Screen Cinema Subtitling System,” [s.d.])

Se é o sistema da Sony, Epson ou Ezra que será eventualmente utilizado em massa, só o tempo dirá, mas há uma interessante ironia nessa história. Nos primórdios do cinema, os filmes mudos eram 'democráticos' enquanto podiam ser apreciados tanto por ouvintes quanto por deficientes auditivos, pois os intertítulos não discriminavam entre os dois. Com o advento dos filmes falados, os deficientes auditivos foram amplamente excluídos; cortados de uma fonte de entretenimento que uma vez desfrutaram. A legendagem inicial era esporádica e indisciplinada, sem diretrizes da indústria sobre o que deveria ou não ser traduzido, (NORNES, M., 2007, p. 25) então, especialmente no caso da legendagem em japonês, algumas cenas foram deixadas parcialmente traduzidas ou não traduzidas. Agora, com o desenvolvimento de sistemas para ajudar os deficientes auditivos, essa abordagem de retorno às raízes pode anunciar o futuro da legendagem para todos.

5.2 CLOSED CAPTION NO BRASIL

A Legislação pró-acessibilidade

Considerando as três dimensões básicas para a implantação de uma política públicas:

1) A segurança jurídica — com as leis que a apoiem; 2) Os serviços e apoios do Estado para que ela ocorra; e por fim; 3) O direito a informações sobre ela — para que os usuários possam reivindicá-las, a lei que regulamenta as primeiras obrigatoriedades em favor às pessoas com deficiência, só foi sancionada após muita luta dos órgãos e instituições de apoio dessas. Através da Lei n.º 10.098 de 19 de dezembro de 2000 no Art. 112 Par. IX:

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

IX — Comunicação: forma de interação dos cidadãos que abrange, entre outras opções, as línguas, inclusive a Língua Brasileira de Sinais (Libras), a visualização de textos, o Braille, o sistema de sinalização ou de comunicação tátil, os caracteres ampliados, os dispositivos multimídia, assim como a linguagem simples, escrita e oral, os sistemas auditivos e os meios de voz digitalizados e os modos, meios e formatos aumentativos e alternativos de comunicação, incluindo as tecnologias da informação e das comunicações. BRASIL, 2000.

Em resumo, esse outorgamento estabelece normas gerais e critérios básicos para a implementação da acessibilidade das pessoas portadoras de deficiência ou com mobilidade reduzida, e dá outras providências. Como a Lei n.º 10.048 de 8 de novembro de 2000, que dá prioridade de atendimento às pessoas com deficiência, a Instrução Normativa n.º 1 de 2 de dezembro de 2005, que regulamenta a utilização dos recursos de acessibilidade para pessoas com deficiência auditiva e visual, e a Portaria n.º 310 de 27 de junho de 2006 (ANATEL, [s.d.]) que estipula recursos de acessibilidade, para pessoas com deficiência, na programação veiculada nos serviços de radiodifusão de sons e imagens e de retransmissão de televisão. A partir desse momento foi possibilitado ao deficiente auditivo e visual maior acesso e inclusão às informações e conteúdo que todos os outros também cidadãos têm. Entretanto, mesmo com essa primeira regulamentação, pouca coisa evoluiu, até que em 2015 com a Lei Brasileira de Inclusão da Pessoa com Deficiência (n.º 13.146 de julho de 2015) que trouxe a obrigatoriedade do uso de tecnologias assistivas em TV aberta. E por conta dessa lei o uso do *Closed Caption*, aumentou significativamente de apenas duas horas diárias de programação com legendagem assistiva para vinte horas diárias, se popularizando como a ferramenta assistiva audiovisual mais aplicada no Brasil (*Closed caption: saiba como funcionam as legendas automáticas nas TVs, s.d.*).

Tecnologias Assistivas presentes no Brasil

Por muito tempo tivemos apenas o *Closed Caption* como ferramenta assistiva, principalmente pelo fato do desenvolvimento de novas tecnologias que facilitariam ainda mais o surdo a consumir o conteúdo disponibilizado com essas tecnologias, como é o caso da *Áudio Descrição*.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

Sobre a áudio descrição pode-se afirmar que:

É uma modalidade de tradução audiovisual, de natureza intersemiótica, que visa tornar uma produção audiovisual acessível às pessoas com deficiência visual. Trata-se de uma locução adicional roteirizada que descreve as ações, a linguagem corporal, os estados emocionais, a ambientação, os figurinos e a caracterização dos personagens. (NEVES, 2016, p.10)

Essa ferramenta é uma tecnologia auditiva direcionada apenas para deficientes visuais, fazendo parte do pacote das tecnologias regulamentadas pela Lei Brasileira de Inclusão da Pessoa com Deficiência. Ela é uma tecnologia obrigatória, determinada a veiculação de ao menos seis horas semanais, todavia já vigora no atual cronograma da ementa regulamentar a veiculação de vinte horas semanais de programação assistida pela ferramenta de áudio descrição (Goodbros Empatia e Acessibilidade e Inclusão, [s.d.]).

Seguindo com as tecnologias assistivas, a mais recente delas, porém menos utilizada por não ser obrigatória na programação comum, apenas em programas de propaganda político-partidária e eleitoral, além de campanhas institucionais e informativos de utilidade pública, é a Janela de Interpretação de Linguagem de Sinais, a qual aqui no Brasil é a LIBRAS, Língua Brasileira de Sinais. Dialeto esse oficializada por meio da Lei 10.436 de 2 de abril de 2002, como a segundo idioma do Brasil.

A Janela de Interpretação nada mais é do que:

O espaço destinado à tradução entre uma língua de sinais e outra língua oral ou entre duas línguas de sinais, feita por Tradutor e Intérprete de Língua de Sinais (TILS), na qual o conteúdo de uma produção audiovisual é traduzido num quadro reservado, preferencialmente, no canto inferior esquerdo da tela, exibido simultaneamente à programação. (NEVES, 2016, p.10)

Aqui temos um ponto muito relevante, que evidencia a importância da Janela de Interpretação para a população surda. O Censo Brasil 2010 (“IBGE Censo 2010,” [s.d.]), não trouxe com exatidão informações sobre a população surda brasileira, como quantos deles são usuários de LIBRAS e seus níveis de alfabetização, apenas um recorte de quantas pessoas são alfabetizadas por região dividido em faixas etárias e os números de pessoas surdas em geral, mas não a taxa de alfabetização de pessoas surdas, nem sua segmentação

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

por faixa etária. Entretanto, estima-se, pelos cálculos da pesquisa do Instituto Locomotiva(Raio-X da Surdez no Brasil, 2019), de que ao menos 40% da população surda não seja alfabetizada em português, tornando a LIBRAS sua língua materna oficial. Sendo assim, temos cerca de quatro milhões de pessoas que dependem da linguagem de LIBRAS, para se comunicarem e consumirem qualquer categoria de conteúdo, tornando essencial a Janela de Interpretação, motivo de discussões entre grupos de inclusão social e líderes governamentais para a obrigatoriedade da ferramenta de Janela de Interpretação também na programação cotidiana da TV aberta.

A ferramenta fonte escolhida para esse estudo, é popularmente conhecida como *Closed Caption*, e no meio técnico por LSE, legendagem para surdos e ensurdecidos. Foi criada em 1971 nos Estados Unidos e apresentada na Primeira Conferência Nacional de Televisão para Deficientes Auditivos, desde então não parou mais de ser utilizada, difundida para os países da Europa. A BBC, em 1971, foi a emissora de TV pioneira a exibir a legendagem encapsulada com base em seu tele-texto. Aqui no Brasil a primeira emissora a exibir legendagem no modelo *Closed Caption*, foi a Rede Globo em 2000, sendo o programa do Jô o precursor.

Em questão de definição, podemos dizer que a *Closed Caption*, ou LSE é:

[]É a tradução das falas de uma produção audiovisual em forma de texto escrito, podendo ocorrer entre duas línguas orais, entre uma língua oral e outra de sinais ou dentro da mesma língua. Por ser voltada, prioritariamente, ao público Surdo e Ensurdecido, a identificação de personagens e efeitos sonoros deve ser feita sempre que necessário. (NEVES, 2016, p.10)

O Estenotipista

Os estenotipistas são treinados em métodos taquigráficos que lhes permitem criar transcrições escritas literais usando uma máquina de estenótipo. O trabalho de um estenógrafo é vital principalmente durante depoimentos pré-julgamento, audiências e julgamentos judiciais, e para a produção de legendas em produtos audiovisuais. Eles, também chamados de estenógrafos, são responsáveis por produzir um registro escrito preciso e completo para que os receptores possam entender ao lerem o conteúdo construído.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

Os estenotipistas especializados em legendagem ao vivo para conferências, reuniões de negócios, programas de televisão ou eventos especiais, devem possuir certos requisitos e habilidades como velocidade e precisão e a capacidade de se concentrar, ouvir e prestar atenção aos detalhes, passando seus dias digitando em uma máquina de taquigrafia chamada estenótipo. Eles digitam por sílabas no estenótipo, sendo um método muito mais rápido do que a taquigrafia manual ou a digitação de letras em um teclado de computador. De acordo com Naegeli (Naegeli, [s.d.]), estenotipistas proficientes podem digitar mais rápido que trezentas palavras por minuto, o que é o dobro da velocidade média de fala humana. A velocidade de digitação é essencial para capturar o diálogo quando várias pessoas estão falando.

Estenógrafos também são muitas vezes empregados como repórteres da corte pelos governos locais e estaduais. Normalmente trabalham para os tribunais ou legislaturas para transcrever procedimentos legais. Posições semelhantes estão disponíveis em escritórios de advocacia, empresas e instalações médicas que preferem um registro escrito de conversas e reuniões. O trabalho de um estenotipista também pode incluir legendas ao vivo ou pós-produção para o benefício de telespectadores ou membros da audiência com deficiência auditiva. Os provedores de tradução em tempo real de acesso à comunicação (CART) criam transcrições em tempo real para clientes com deficiência auditiva em ambientes como conferências, aulas na faculdade e consultas médicas.

É uma tarefa bastante estressante porque não há espaço para erros. Um estenógrafo deve ficar focado e alerta ao transcrever e as expectativas do empregador quanto ao desempenho são altas.

O Bureau of Labor Statistics, dos Estados Unidos, antecipa o aumento da demanda pelo trabalho de um estenógrafo. Especificamente, o BLS prevê uma taxa de crescimento de 9% entre 2019 e 2029, mais rápido que o crescimento de 4% de todas as ocupações na economia dos EUA. As perspectivas de emprego serão melhores para aqueles que se formarem em um programa de reportagem do tribunal e para especialistas em legendas ao vivo, como observamos na figura 7 abaixo. (Court Reporters: Occupational

Outlook Handbook:: U.S. Bureau of Labor Statistics, [s.d.]



Figura 7 – Estenotipista criando Closed Caption em teclado personalizado para a função.
Fonte: YouTube — Disponível em: <https://www.youtube.com/watch?v=ZAKkTtsPKOA>

5.3 PYTHON — LINGUAGEM DE PROGRAMAÇÃO

5.3.1 História do Python

Python é uma linguagem de programação de uso geral, amplamente utilizada e reconhecida atualmente, que significa que pode ser usada não apenas para desenvolvimento web, mas para back-end, ciência de dados de aplicação de software e escrita de scripts de sistema. Muitas empresas populares a utilizam em suas pilhas de tecnologia, como Instagram, Google, Spotify e YouTube. Ele foi projetado por Guido van Rossum na década de 1980 e desenvolvido pela Python Software Foundation, sendo que seu principal princípio é enfatizar a legibilidade do código. A ABC, outra linguagem de programação de uso geral desenvolvida na Holanda, teve uma influência muito forte na criação do Python. Rossum ajudou a construir o ABC e expressou suas experiências positivas e negativas com ela. Ele queria criar uma linguagem que tivesse algumas das maiores propriedades do ABC, com sintaxe simples e tipos de dados poderosos. A ideia do Python nasceu então e mais tarde foi lançada em 1991 (History and Philosophy of Python | Python Tutorial | python-course.eu, [s.d.]).

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

Ao contrário de outras linguagens, como Java e C, não são necessárias tantas linhas de código para executar a funcionalidade, que é uma das razões pelas quais é mais fácil de aprender, especialmente para pessoas que estão começando no mundo da programação ou para novos desenvolvedores em geral. Semelhante ao Ruby, a sintaxe do Python é composta principalmente de comandos semelhantes ao inglês, e isso o torna mais eficiente também.

Embora o Python seja uma linguagem fácil de decifrar e leve menos linhas de código em desempenho, na verdade, é bastante lenta quando os arquivos são executados. Pode ser tanto uma linguagem interpretativa quanto compilada, no entanto, é principalmente considerada interpretativa devido à execução de linha por linha durante sua confecção. As linguagens compiladas, por outro lado, se convertem diretamente em código de máquina que o processador pode executar. A razão pela qual isso torna o Python uma linguagem mais lenta é porque suas instruções devem ser entendidas pela CPU do computador antes de serem finalmente executadas. Em outras palavras, o Python verifica se o seu código obedece às suas regras e faz isso linha por linha, adicionando significativamente à execução geral do programa.

Apesar de o tempo de desempenho ser lento, a economia de tempo no processo de desenvolvimento é mais rápida em comparação com outras linguagens. O Python é realmente uma ótima linguagem para desenvolvimento web e de jogos, além de aplicativos de desktop e ciência/visualização de dados, fazendo dele hoje a linguagem de programação mais utilizada no mundo (TIOBE Index, 2022).

5.3.2 Python na Transcrição de Áudio em Texto

Antes de chegarmos a questão da transcrição áudio em Python, falemos um pouco sobre como o reconhecimento de fala funciona. Ele tem suas raízes em pesquisas feitas no Bell Labs no início da década de 1950. Os primeiros sistemas eram limitados a um único falante e tinham vocabulários limitados de cerca de uma dúzia de palavras (MOSKVITCH, 2017). Os sistemas modernos de reconhecimento de voz percorreram um longo caminho desde os seus homólogos antigos, pois agora eles podem reconhecer a fala de vários

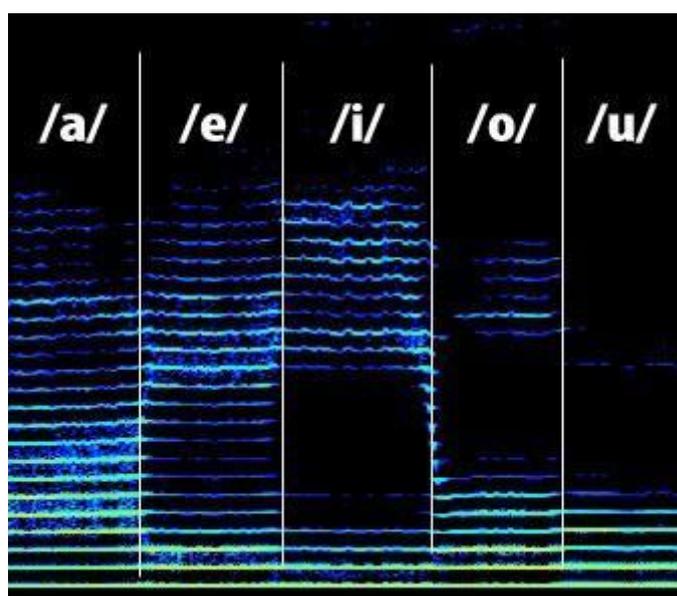
Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

falantes e têm vocabulários enormes em vários idiomas.

O primeiro componente do reconhecimento de fala é, obviamente, a fala. Esta deve ser convertida de som físico para um sinal elétrico com um microfone e, em seguida, para dados digitais com um conversor analógico-digital. Uma vez digitalizados, vários modelos podem ser usados para transcrever o áudio para texto.

A maioria dos sistemas modernos de reconhecimento de fala e transcrição de áudio se baseia no que é conhecido como Hidden Markov Model (HMM). Essa abordagem funciona com a suposição de que um sinal de fala, quando visto em uma escala de tempo suficientemente curta (alguns milissegundos, por exemplo), pode ser razoavelmente aproximado como um processo estacionário – isto é, com propriedades estatísticas que não mudam ao longo do tempo. (EDDY, 2004, p.1315)

Em um HMM típico, o sinal de fala é dividido em fragmentos de vários milissegundos. O espectro de potência de cada fragmento, que é essencialmente um gráfico da potência do sinal em função da frequência, é mapeado para um vetor de números reais conhecido como coeficientes cepstrais. A dimensão desse vetor é geralmente pequena – às vezes tão baixa quanto dez, embora sistemas mais precisos possam ter dimensão 32 ou mais. A saída final do HMM é uma sequência desses vetores (EDDY, 2004, p. 1316).



Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

Figura 8 - Espectro de potência de fragmentos das vogais quando faladas. Fonte: Estúdio de Voz - Mauro Fiuza. 2015. Disponível em: <http://www.estudiodevoz.com.br/2015/07/formantes-ressonancias-sintonizacao-de_29.html>.

Atualmente, para decodificar a fala em texto, grupos de vetores são combinados com um ou mais fonemas – que nada mais é do que uma unidade fundamental da fala. Esse cálculo requer treinamento, pois o som de um fonema varia de um locutor para outro e até de um enunciado para outro do mesmo locutor. Um algoritmo especial é então aplicado para determinar a palavra (ou palavras) mais provável que produz a sequência de fonemas dada.

Todo esse processo pode ser computacionalmente caro, entretanto com o avanço da tecnologia, adquirimos possibilidades de desenvolver novas soluções. Em muitos sistemas modernos de reconhecimento de fala, as redes neurais são usadas para simplificar o sinal de fala usando técnicas para transformação de recursos e redução de dimensionalidade antes do reconhecimento HMM. Os detectores de atividade de voz (VADs) também são utilizados para reduzir um sinal de áudio apenas para as partes que provavelmente contêm fala. Isso evita que o reconhecedor perca tempo analisando partes desnecessárias do sinal.

5.3.3 Bibliotecas¹ de Transcrição de Áudio em Texto que Utilizam Machine Learning

- **SpeechRecognition**

A biblioteca SpeechRecognition atua como um empacotador para várias APIs de fala populares e, portanto, é extremamente flexível. Uma delas, a Google Web Speech APL, que oferece suporte a uma chave de API padrão que é codificada na biblioteca SpeechRecognition. Isso significa que você pode utilizá-lo sem precisar se inscrever em um serviço pago.

A flexibilidade e a facilidade de uso deste pacote o tornam uma excelente opção para qualquer projeto Python e por esse motivo ela foi a escolhida para comandar esse projeto.

- **assemblyai**

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

É uma biblioteca recente, com uma interface de fácil uso e recursos como detecção de alto-falante, de tópico, de parágrafo e de pontuação e caixa automatizadas, entre outros. Sua alta precisão e extensa lista de recursos, como Diarização de alto-falante e Análise de sentimento, a torna uma boa opção para desenvolvedores que procuram uma API de Speech-to-Text. Esta API também suporta praticamente todos os formatos de arquivos de áudio e vídeo prontos para uso para facilitar a transcrição. Ela seria a opção perfeita se não tivesse custo, além de oferecer suporte apenas à transcrição em inglês, limitando seu uso a países de língua inglesa e possui SDKs limitados disponíveis.

- **Pocketsphinx**

O PocketSphinx, é um dos mecanismos de reconhecimento de fala contínuo e independente de alto-falante da Carnegie Mellon University. Ele foi desenvolvido para melhor compatibilidade com sistemas de transcrição de áudio que possuem hardwares externos, como placas de captura de áudio, moduladores e equalizadores. Por esse motivo ele não tem total funcionamento online, dependendo de ferramentas físicas para seu melhor desempenho.

- **wit**

Este é uma biblioteca especialmente desenvolvida com melhor compatibilidade com produtos de Realidade Aumentada e Realidade Virtual. Possui integração nativa com os óculos de Realidade Virtual, sendo a opção perfeita para a transcrição de áudio no Metaverso e em aplicações de RA e RV.

5.3.4 História do Machine Learning

O aprendizado de máquina (Machine Learning) é uma ferramenta importante para o objetivo de alavancar tecnologias em torno da inteligência artificial. Por causa de suas habilidades de aprendizado e tomada de decisão, o aprendizado de máquina é frequentemente chamado de IA, embora, na realidade, seja uma subdivisão dela, que até o final da década de 1970, fazia

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

parte da evolução da IA. Então, ele se ramificou para evoluir por conta própria. O aprendizado de máquina tornou-se uma ferramenta de resposta muito importante para computação em nuvem e comércio eletrônico e está sendo usado em uma variedade de tecnologias de ponta.

O aprendizado de máquina é um aspecto necessário dos negócios e da pesquisa moderna para muitas organizações hoje. Ele usa algoritmos e modelos de redes neurais para auxiliar os sistemas de computador a melhorar progressivamente seu desempenho. Os algoritmos de aprendizado de máquina criam automaticamente um modelo matemático usando dados de amostra – também conhecidos como “dados de treinamento” – para tomar decisões sem serem especificamente programados para tomar essas decisões. Em parte ele é baseado no modelo de interação das células cerebrais. Foi criado em 1949 por Donald Hebb em um livro intitulado “*The Organization of Behavior*”, no qual se apresenta as teorias de Hebb sobre excitação de neurônios e comunicação entre neurônios.

Hebb (1949) escreveu: “Quando uma célula repetidamente ajuda a disparar outra, o axônio da primeira célula desenvolve botões sinápticos (ou aumenta-os se já existirem) em contato com a soma da segunda célula”. Traduzindo os conceitos de Hebb para redes neurais artificiais e neurônios artificiais, seu modelo pode ser descrito como uma forma de alterar as relações entre neurônios artificiais (também chamados de nós) e as mudanças em neurônios individuais. A relação entre dois neurônios/nós se fortalece se os dois neurônios/nós forem ativados ao mesmo tempo e enfraquece se forem ativados separadamente. A palavra “peso” é usada para descrever essas relações, e nossos neurônios que tendem a ser positivos ou negativos são descritos como tendo pesos positivos fortes. Aqueles nós que tendem a ter pesos opostos desenvolvem fortes pesos negativos (por exemplo: $1 \times 1 = 1$, $-1 \times -1 = 1$, $-1 \times 1 = -1$)

Arthur Samuel da IBM desenvolveu um programa de computador (11.2 Samuel’s Checkers Player, [s.d.]) para jogar damas na década de 1950. Como o programa tinha uma quantidade muito pequena de memória de computador disponível, Samuel iniciou o que é chamado de poda alfa-beta (“Step by Step:

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

Alpha Beta Pruning,” [s.d.]). Seu projeto incluía uma função de pontuação usando as posições das peças no tabuleiro. A função de pontuação tentou medir as chances de cada lado ganhar. O programa escolhe seu próximo movimento usando uma estratégia minimax, que eventualmente evoluiu para o algoritmo minimax (HINTZE, [s.d.]).

Samuel também projetou vários mecanismos que permitiram que seu programa se tornasse melhor. No que Samuel chamou de aprendizado mecânico, seu programa registrou/lembrou todas as posições que já havia visto e combinou isso com os valores da função de recompensa. Arthur Samuel surgiu com a frase “aprendizagem de máquina” em 1952.

Em 1957, Frank Rosenblatt – no Cornell Aeronautical Laboratory – combinou o modelo de interação de células cerebrais de Donald Hebb com os esforços de aprendizado de máquina de Arthur Samuel e criou o perceptron. Este inicialmente foi planejado como uma máquina, não como um programa. O software, originalmente projetado para o IBM 704, foi instalado em uma máquina customizada chamada de Marca 1 perceptron (JUSTIN, B., 2017), que havia sido construída para reconhecimento de imagem. Isso tornou o software e os algoritmos transferíveis e disponíveis para outras máquinas.

Descrito como o primeiro neurocomputador de sucesso, o perceptron Mark I desenvolveu alguns problemas com expectativas quebradas. Embora o perceptron parecesse promissor, ele não conseguia reconhecer muitos tipos de padrões visuais (como rostos), causando frustração e paralisando a pesquisa de redes neurais. Levaria vários anos antes que as frustrações de investidores e agências de financiamento desaparecessem. A pesquisa de aprendizado de máquina lutou até um ressurgimento durante a década de 1990.

Trinta anos antes, em 1967, foi concebido o algoritmo do vizinho mais próximo, que foi o início do reconhecimento básico de padrões. Esse algoritmo foi usado para mapear rotas e, um dos primeiros, para encontrar uma solução para o problema de o vendedor ambulante achar a rota mais eficiente. Usando-o, um vendedor entra em uma cidade selecionada e

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

repetidamente faz o programa visitar as cidades mais próximas até que todas tenham sido visitadas. Marcello Pelillo recebeu crédito por inventar a “regra do vizinho mais próximo”. Ele, por sua vez, credita o famoso artigo de Thomas Cover em 1967 (COVER, 1967).

Na década de 1960, a descoberta e o uso de multicamadas abriram um novo caminho na pesquisa de redes neurais. Descobriu-se que fornecer e usar duas ou mais camadas no perceptron oferecia significativamente mais poder de processamento do que um perceptron usando uma camada. Outras versões de redes neurais foram criadas depois, mas este abriu a porta para “camadas” nas redes, e a variedade destas continua a se expandir. O uso de várias camadas levou a redes neurais de feed forward retropropagação (NIELSEN, 2019).

A retropropagação desenvolvida na década de 1970, permite que uma rede ajuste suas camadas ocultas de neurônios/nós para se adaptar a novas situações. Ele descreve “a propagação de erros para trás”, com um erro sendo processado na saída e então distribuído para trás pelas camadas da rede para fins de aprendizado. A retropropagação agora está sendo usada para treinar redes neurais profundas.

A rede neural artificial (ANN) tem camadas ocultas que são usadas para responder a tarefas mais complicadas do que os *perceptrons* anteriores. As RNAs são uma ferramenta primária usada para aprendizado de máquina. As redes neurais usam camadas de entrada e saída e, normalmente, incluem uma camada oculta (ou camadas) projetadas para transformar a entrada em dados que podem ser usados pela camada de saída. As camadas ocultas são excelentes para encontrar padrões muito complexos para um programador humano detectar, o que significa que um humano não poderia encontrar o padrão e depois ensinar o dispositivo a reconhecê-lo.

No final da década de 1970 e início da década de 1980, a pesquisa se concentrou no uso de abordagens lógicas baseadas no conhecimento, em vez de algoritmos. Além disso, a pesquisa de redes neurais foi abandonada por pesquisadores de ciência da computação e IA. Isso causou uma cisma entre inteligência artificial e aprendizado de máquina. Até então, o aprendizado de

máquina era usado como um programa de treinamento para IA.

A indústria de aprendizado de máquina, que incluía muitos pesquisadores e técnicos, foi reorganizada em um campo separado e lutou por quase uma década (KURENKOV, 2020). O objetivo da indústria mudou de treinamento para inteligência artificial para resolver problemas práticos em termos de prestação de serviços. Seu foco passou de abordagens herdadas da pesquisa de IA para métodos e táticas usadas em teoria de probabilidade e estatística. Durante esse período, a indústria de ML manteve seu foco em redes neurais e floresceu na década de 1990. A maior parte desse sucesso foi resultado do crescimento da Internet, beneficiando-se da disponibilidade cada vez maior de dados digitais e da capacidade de compartilhar seus serviços por meio da Internet.

O “*Boosting*” algoritmos (ANALYTICS VIDHYA, 2016) foi necessário para a evolução do aprendizado de máquina, eles são usados para reduzir o viés durante o aprendizado supervisionado e incluem algoritmos de ML que transformam alunos fracos em fortes. O conceito de boosting foi apresentado pela primeira vez em um artigo de 1990 intitulado “The Strength of Weak Learnability”, de Robert Schapire. Ele afirma (SCHAPIRE, 1989) que: “um conjunto de alunos fracos pode criar um único aluno forte”. Alunos fracos são definidos como classificadores que são apenas ligeiramente correlacionados com a classificação verdadeira (ainda melhor do que adivinhação aleatória). Por outro lado, um aprendiz forte é facilmente qualificado e bem alinhado com a verdadeira classificação.

A maioria dos algoritmos de boosting são compostos de classificadores fracos de aprendizado repetitivo, que então se somam a um classificador forte final. Depois de adicionados, eles são normalmente ponderados de forma a avaliar a precisão dos alunos fracos. Em seguida, os pesos dos dados são “re-ponderados”. Os dados de entrada classificados incorretamente ganham um peso maior, enquanto os dados classificados corretamente perdem peso. Este ambiente permite que futuros alunos fracos se concentrem mais extensivamente em alunos fracos anteriores que foram classificados incorretamente.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

A diferença básica entre os vários tipos de algoritmos de reforço é “a técnica” usada na ponderação dos pontos de dados de treinamento. O *AdaBoost* é um algoritmo de aprendizado de máquina popular e historicamente significativo, sendo o primeiro algoritmo capaz de trabalhar com alunos fracos (BROWNLEE, 2016). Algoritmos mais recentes incluem *BrownBoost*, *LPBoost*, *MadaBoost*, *TotalBoost*, *xgboost* e *LogitBoost*. Muitos algoritmos de reforço funcionam na estrutura *AnyBoost*.

Atualmente, grande parte do treinamento de reconhecimento de fala está sendo feito por uma técnica de Deep Learning chamada memória de longo prazo (LSTM), um modelo de rede neural descrito por Jürgen Schmidhuber e Sepp Hochreiter em 1997. LSTM pode aprender tarefas que exigem memória de eventos que ocorreram milhares de etapas discretas antes, o que é muito importante para a fala. (MAYO, 2018). Segundo Mayo (2018), por volta do ano de 2007, a memória de curto prazo começou a superar os programas de reconhecimento de fala mais tradicionais. Em 2015, o programa de reconhecimento de fala do Google supostamente teve um salto significativo de desempenho de 49% usando um LSTM treinado pelo CTC.

Em 2006, o Instituto Nacional de Padrões e Tecnologia realizou o “*Face Recognition Grand Challenge*” – que avaliou os algoritmos de reconhecimento facial populares da época. Varreduras de rosto em 3D, imagens de íris e imagens de rosto de alta resolução foram testadas. Suas descobertas sugeriram que os novos algoritmos eram dez vezes mais precisos do que os de reconhecimento facial de 2002 e cem vezes mais precisos do que os de 1995. Alguns dos algoritmos foram capazes de superar os participantes humanos no reconhecimento de rostos e identificar exclusivamente gêmeos idênticos (“Face Recognition Grand Challenge (FRGC)” [s.d.]).

Em 2012, o *XLab* do Google desenvolveu um algoritmo de ML que pode navegar e encontrar de forma autônoma vídeos contendo gatos. Em 2014, o Facebook desenvolveu o *DeepFace*, um algoritmo capaz de reconhecer ou verificar indivíduos em fotografias com a mesma precisão dos humanos.

O aprendizado de máquina agora é responsável por alguns dos avanços mais significativos da tecnologia. Está sendo usado para a nova indústria de

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

veículos autônomos, e para explorar a galáxia, pois ajuda na identificação de exoplanetas. Recentemente, o aprendizado de máquina foi definido pela Universidade de Stanford como “a ciência de fazer os computadores agirem sem serem explicitamente programados”. O aprendizado de máquina gerou uma nova gama de conceitos e tecnologias, incluindo aprendizado supervisionado e não supervisionado, novos algoritmos para robôs, Internet das Coisas, ferramentas de análise, chatbots, entre outros.

Os modelos de aprendizado de máquina tornaram-se bastante adaptativos no aprendizado contínuo, o que os torna cada vez mais precisos quanto mais tempo operam. Algoritmos de ML combinados com novas tecnologias de computação promovem a escalabilidade e melhoram a eficiência. Combinado com a análise de negócios, o aprendizado de máquina pode resolver uma variedade de complexidades organizacionais. Modelos modernos de ML podem ser usados para fazer previsões que vão desde surtos de doenças até aumento e queda de estoques. Atualmente, o Google está experimentando o aprendizado de máquina usando uma abordagem chamada ajuste fino de instruções. O objetivo é treinar um modelo de ML para resolver problemas de processamento de linguagem natural de maneira generalizada. O processo treina o modelo para resolver uma ampla gama de problemas, em vez de apenas um tipo de problema (MONTTI, 2021).

5.3.5 Machine Learning na Acessibilidade

Com o avanço da tecnologia digital, muitas soluções em auxílio às ferramentas de acessibilidade foram surgindo, com base nos aportes da inteligência artificial, que é

[].um processo que envolve o agrupamento de várias tecnologias, como redes neurais artificiais, algoritmos, sistemas de aprendizado, entre outros que conseguem simular capacidades humanas ligadas à inteligência. Por exemplo, o raciocínio, a percepção de ambiente e a habilidade de análise para a tomada de decisão.” (O que é Inteligência artificial? Como funciona, exemplos e aplicações, 2019).

Através de modelos de processamento de linguagem neural, o *Machine Learning*, também conhecido como Aprendizado de Máquina, apoiado pelas

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

linguagens de programação de bibliotecas de códigos pré-construídas, por exemplo, a biblioteca *Speech Recognition* utilizada como alicerce tecnológico nesta pesquisa, pelo fato de ter maior facilidade de programação e maior acurácia em relação às outras, e quando comparada em testes práticos com software utilizado pelo *Closed Caption*, o atraso é muito menor e possui assertividade linguística semelhante.

Outra ferramenta muito utilizada atualmente que faz o uso constante de modelos de processamento de linguagem neural e módulos de inteligência artificial é o *Hand Talk*, que se constitui de vários módulos, que fazem a leitura do conteúdo em texto e traduzem automaticamente para interpretação de LIBRAS, expressando os sinais através de uma janela de interpretação. Isso é possível principalmente por conta da combinação do módulo de inteligência artificial, com o método *Machine Learning*, que faz a leitura inicial dos dados, nesse caso os textos, e automatiza os modelos para o módulo de IA. Essa performance conjunta, executada repetitiva e constantemente, ocasiona no aprendizado e aperfeiçoamento da leitura dos dados, melhorando o desempenho da ferramenta a cada utilização, transcrevendo com maior efetividade os textos.

Para este estudo foi utilizada a biblioteca de códigos *Speech Recognition* para a interpretação dos dados vindos de arquivos de áudio e transcrevê-los em forma texto em uma tela. A *Speech Recognition* é muito mais simples do que os sistemas atuais utilizados, tanto na sua implementação em produtos audiovisuais, quanto em seu uso individual em um computador pessoal, e, em simultâneo, fornece uma maior qualidade de reconhecimento da fala na presença de ruído estranho.

O desenvolvimento desta biblioteca não usa modelos acústicos tradicionais e o conceito de fonemas²; em vez disso, usa um sistema de aprendizado de máquina com base em rede neural bem otimizada, eliminando a necessidade de desenvolver componentes separados para modelar vários desvios, como ruído, eco e características de fala.

O outro lado dessa abordagem é que, para obter reconhecimento e treinamento de alta qualidade de uma rede neural, o motor *Speech*

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

*Recognition*³ requer uma abundância de dados heterogêneos dissertados em condições reais por vozes diferentes e na presença de ruído natural.

A escolha da utilização dessa ferramenta deu-se pela dupla funcionalidade: além de ser um módulo interpretativo dos dados provenientes dos arquivos de áudio, ele possui também a capacidade de aprendizado de máquina, sem a necessidade de um módulo externo para isso. Sendo assim, existe uma economia significativa no custo do desenvolvimento e implantação da ferramenta, muito próximo de zero. Contudo, há vários outros softwares para descrever o ato de falar, em sua totalidade, eles captam as vibrações orais e elas são convertidas por meio de um conversor analógico-digital (ADC) em dados digitais, para que o computador possa entender. Para isso, ele coleta ou digitaliza o som por meio de medições precisas de reverberação de onda em intervalos de frequência — é a análise dimensional. Logo após, um algoritmo filtra o som digitalizado para remover ruídos indesejados e, às vezes, separá-lo em diferentes faixas de frequência (sendo o comprimento que as ondas sonoras têm, percebidas pelos seres humanos como diferenças de afinação).

Esses softwares são eficientes por conseguirem ajustar o algoritmo à frequência do som a um nível de volume constante. As pessoas nem sempre falam na mesma velocidade, por isso o som também passa por um ajuste fino para corresponder à velocidade das amostras de som do modelo já armazenadas na memória do sistema — sendo as bibliotecas usadas e a inteligência artificial.

5.4 SOFTWARES TRANSCRITORES DE ÁUDIO EM TEXTO

Hoje, no mercado de transcritores de texto temos inúmeras opções, cada um com sua vantagem competitiva perante os concorrentes, uns com maior acurácia, e outros com mais idiomas para transcrição, outros ainda com maior tempo de duração de áudio. Para esse estudo, foram selecionadas as mais compatíveis e semelhantes com o protótipo desenvolvido.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

5.4.1 Happy Scribe

Fundada em 2017, essa empresa irlandesa com sede em Dublin, tem como cerne do seu negócio-produto de transcrição de áudio em texto e legendagem por duas formas, automática e especializada, sendo que esta última a transcrição em feito por um especialista humano, nativo da língua ao qual se deseja efetuar a transcrição. Em seu software de transcrição automática, a empresa promete cerca de 85% de acurácia, entretanto não faz transcrição em tempo real, onde o usuário deve subir o arquivo de áudio na plataforma online deles para ser efetuada a conversão. Os primeiros dez minutos de transcrição são grátis, oferecidos como teste da ferramenta, após o fim desse teste a plataforma cobra por volta de R\$1,20 a cada minuto de áudio enviado para a transcrição.

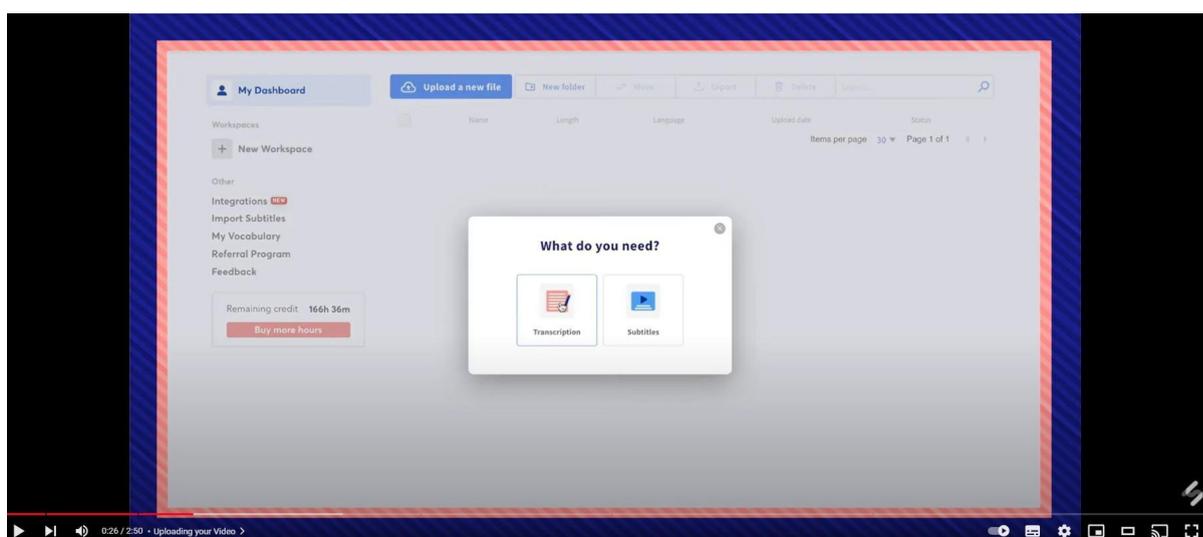


Figura 9 - Vídeo de demonstração de como utilizar a ferramenta Happy Scribe.
Fonte: YouTube. Disponível em: https://www.youtube.com/watch?v=_1UXqtclh30

5.4.2 Rev

Estabelecida em 2010 por cinco colegas de quarto do MIT, a Rev pode ser considerada uma das propulsoras do mercado de transcrição de áudio. Basicamente, todos os outros softwares possuem semelhanças em suas ferramentas com o transcritor automático da Rev. Ela é amplamente reconhecida como provedora de serviços de fala para texto no mundo todo, Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

oferecendo velocidade, precisão, consistência e acessibilidade via seus produtos. Seus serviços de fala para texto são alimentados por inteligência artificial e aperfeiçoados por uma comunidade de freelancers qualificados. Em seu transcritor automático, é garantido cerca de 80% de acurácia, e assim como os outros avaliados nessa lista, não efetua transcrição em tempo real, com o formato de que o usuário precisa subir o arquivo de áudio na plataforma para ser efetuada a conversão. O teste grátis permite ao usuário 45 minutos de transcrição, e quando finalizado a plataforma cobra por volta de R\$1,31 por minuto de áudio enviado para a transcrição.

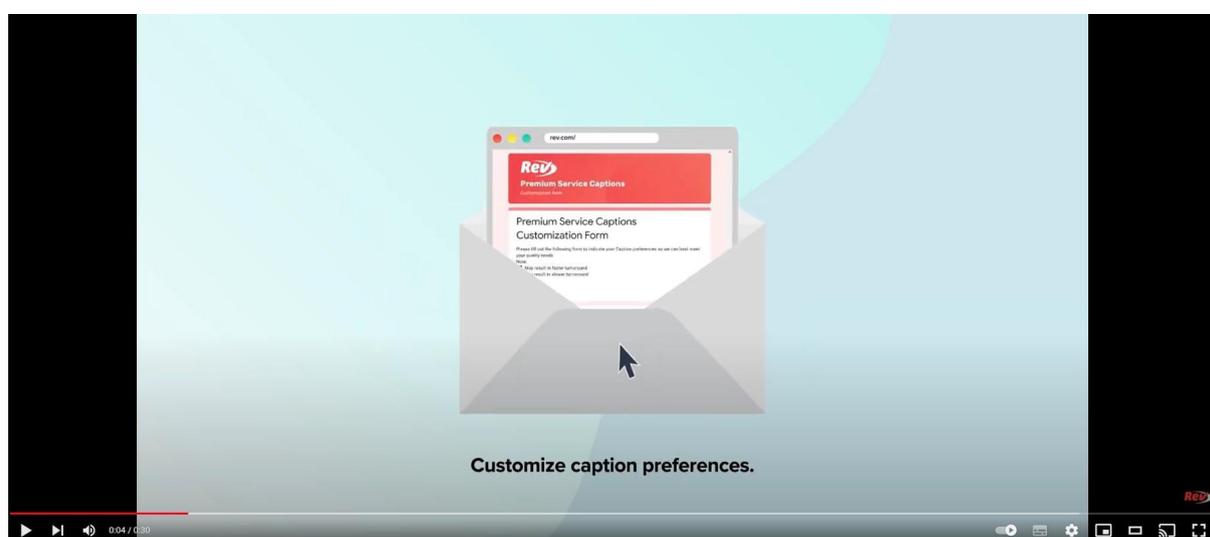


Figura 10 — Vídeo de demonstração de como utilizar a ferramenta Rev. Fonte: YouTube. Disponível em: https://www.youtube.com/watch?v=_1UXqtclh30

5.4.3 Amberscript

Se formos criar uma lista dos melhores e piores, a Amberscript levaria com folga o título de pior ferramenta. Sua baixa gama de idiomas para transcrição (39 idiomas), a acuracidade mediana (85%) e o comprimento de áudio enviado medíocre de sessenta minutos não seriam suficientes para colocá-la nesta posição, se não fosse pelo valor astronômico cobrado pelo minuto de transcrição, que é de cerca de R\$52,40. Sendo 40% mais caro do que segundo colocado nesta ingrata lista dos piores. O que deixa claro que a fonte de receita da companhia vem através de seus outros produtos, como a transcrição manual especializada, semelhante às empresas anteriores.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

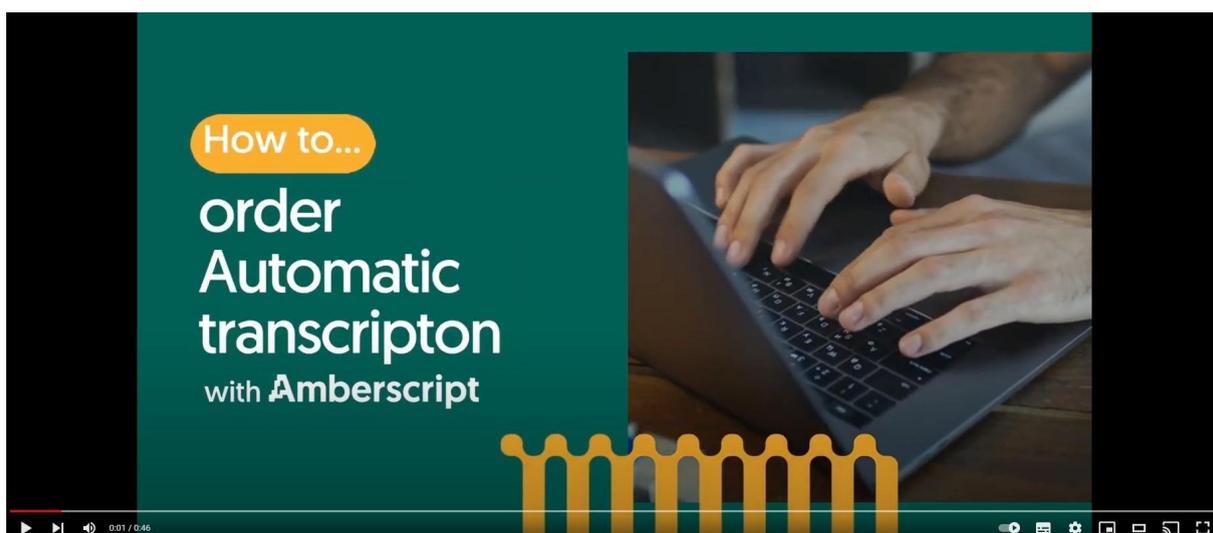
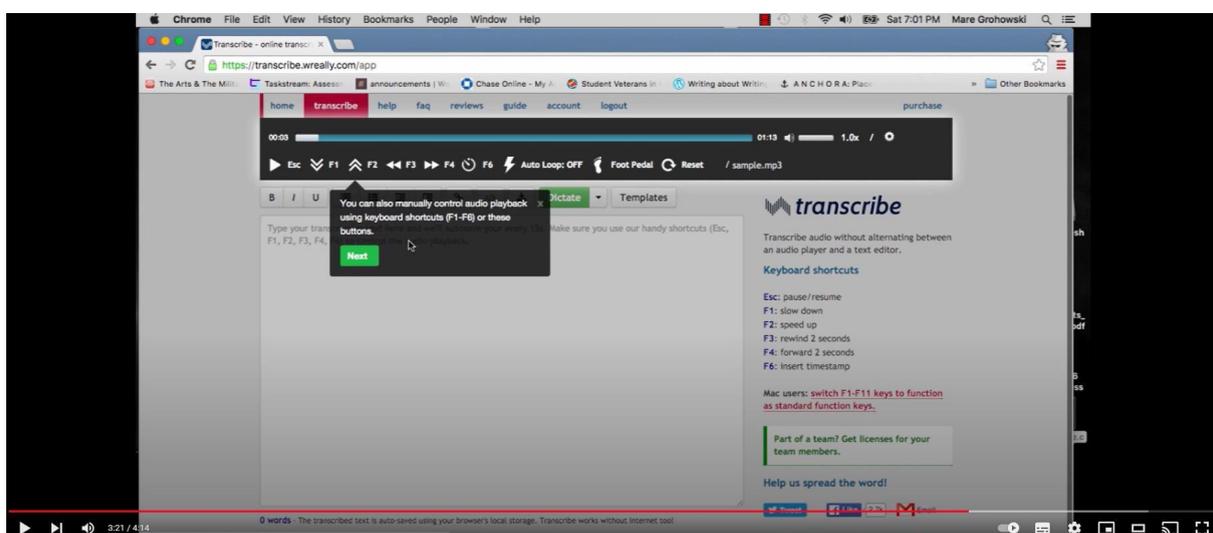


Figura 11 — Vídeo de demonstração de como utilizar a ferramenta Amberscript. Fonte: YouTube. Disponível em: https://www.youtube.com/watch?v=_1UXqtch30

5.4.4 Transcribe

Sendo um produto secundário de uma empresa americana desenvolvedora de softwares, a Wreally, a ferramenta Transcribe é bastante semelhante às anteriores, porém, um pouco mais simples no quesito design e funcionalidades.

Em sua plataforma, que possui seu software de transcrição automática, é assegurado 87% de acurácia, contudo ela não faz transcrição em tempo real. O teste é de apenas um minuto de transcrição grátis, cobrando por volta de R\$0,52 a cada minuto de áudio enviado para a transcrição, sendo uma das ferramentas mais baratas de se utilizar.



Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

Figura 12 — Vídeo de demonstração de como utilizar a ferramenta Transcribe. Fonte: YouTube.

Disponível em: https://www.youtube.com/watch?v=_1UXqtclh30

5.4.5 Trint

Empresa inglesa com sede em Londres, a Trint, foi fundada em 2014 pelo vencedor do Emmy, repórter e correspondente internacional Jeff Kofman para sanar uma dor que ele tinha nos momentos de efetuar a transcrição dos áudios de suas chamadas. Em sua carreira de trinta anos na transmissão, Jeff correspondeu jornalisticamente em mais de quarenta países, cobrindo muitas das maiores histórias de nossa atual geração. Ele queria saber por que ainda estava transcrevendo entrevistas, discursos e coletivas de imprensa no século XXI da mesma forma que os jornalistas faziam na década de 1970: manualmente.

Jeff se perguntou se havia uma maneira de aproveitar a saída cada vez mais precisa da inteligência artificial (IA) e da fala para texto para fazer o trabalho pesado da transcrição. Após colaborar com uma pequena equipe de desenvolvedores, eles inventaram o Trint. Assim como as demais, ela possui uma acurácia de transcrição de cerca de 87%, mas se destaca com seu teste grátis, possibilitando que o usuário faça a transcrição de 540 minutos de áudio, sendo o maior teste dentre as ferramentas do estudo. Após esse teste a empresa segue cobrando apenas R\$0,26 por minuto, sendo também a mais barata, atingindo assim um alto custo-benefício se comparada com as outras.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

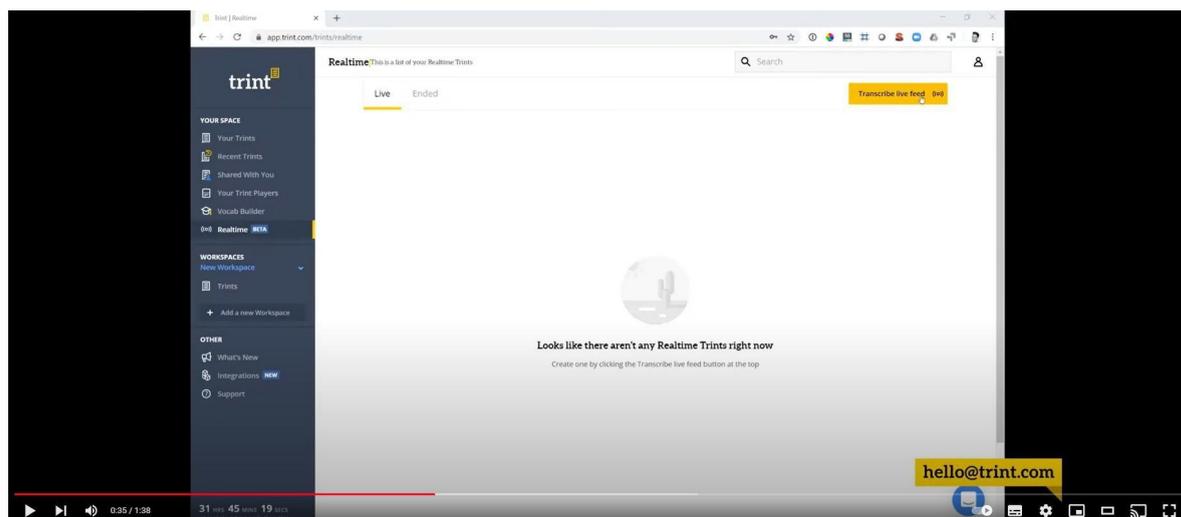


Figura 13 — Vídeo de demonstração de como utilizar a ferramenta Trint. Fonte: YouTube. Disponível em: https://www.youtube.com/watch?v=_1UXqtclh30

5.4.6 Temi

Uma das melhores ferramentas se considerarmos todos os indicadores avaliados. Com a segunda melhor acurácia (92%) a Temi, companhia americana, fundada em 2017, alia qualidade de transcrição com custo-benefício (R\$1,31 por minuto transcrito) para se posicionar como uma excelente alternativa no mercado de transcrição, não fosse sua disponibilidade única de idioma (inglês) e assim como as outras não possui recurso de efetuar transcrição em tempo real. Ela possui um API de integração com outros softwares, possibilitando sua integração com programas de armazenamento e e-mail para transcrever arquivos de áudio contidos nessas ferramentas.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

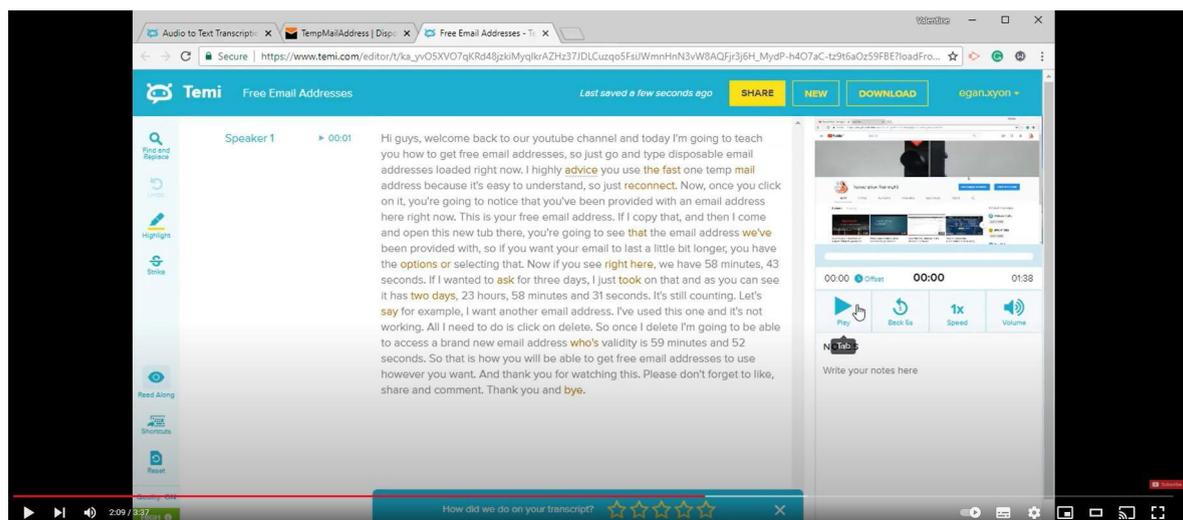


Figura 14 — Vídeo de demonstração de como utilizar a ferramenta Temi. Fonte: YouTube. Disponível em: https://www.youtube.com/watch?v=_1UXqtclh30

5.4.7 Descript

A Descript vende qualidade, esse é o principal fator da empresa não ter o melhor software dessa lista. Possui a melhor acurácia dentre todas as ferramentas (95%) e um excelente período de testes (180 minutos), entretanto, quando pontuamos o quesito custo a ferramenta é a terceira mais cara, além de também só possuir transcrição para um idioma atualmente (inglês). A companhia americana fundada em 2017, tem potencial para ir muito mais longe, necessita apenas de uma expansão no leque de idiomas e um preço mais competitivo.

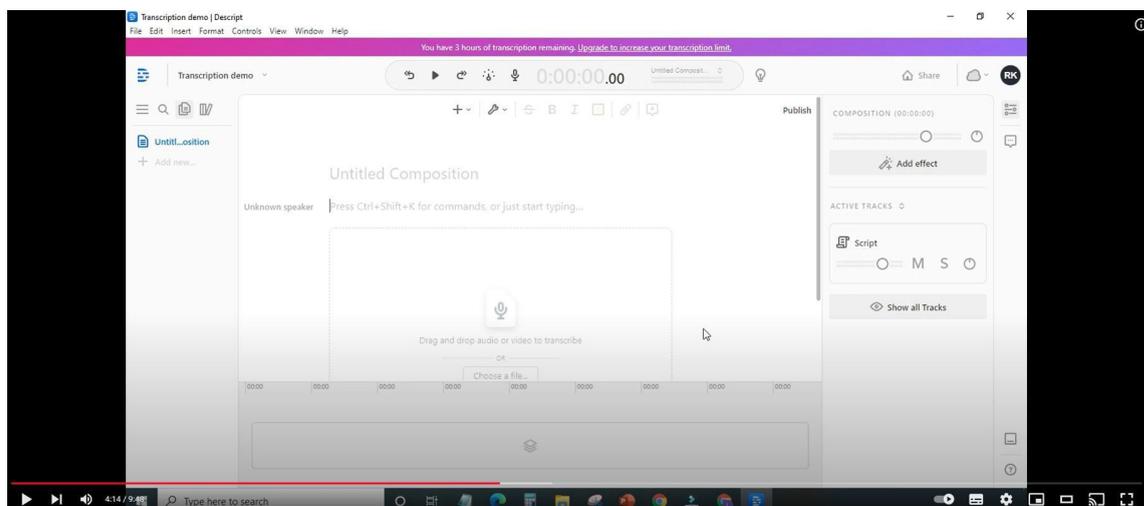


Figura 15 — Vídeo de demonstração de como utilizar a ferramenta Descript. Fonte: YouTube.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

Disponível em: https://www.youtube.com/watch?v=_1UXqtclh30

5.4.8 Sonix

Idêntica a Descript, a Sonix também vende qualidade. Possui a melhor taxa de acurácia (95%), um tempo de teste razoável (30 minutos), uma boa gama de idiomas para transcrição (30 idiomas), mas peca e muito no quesito custo, pois é o maior dentre todos os softwares aqui listados, com cerca de R\$31,44, a avaliação da Sonix tende a ser bastante comprometida, e apenas por esse fato faz com que em uma classificação geral ela ocuparia uma posição bem abaixo do que estaria se estivesse em linha com os valores cobrados pelo mercado.

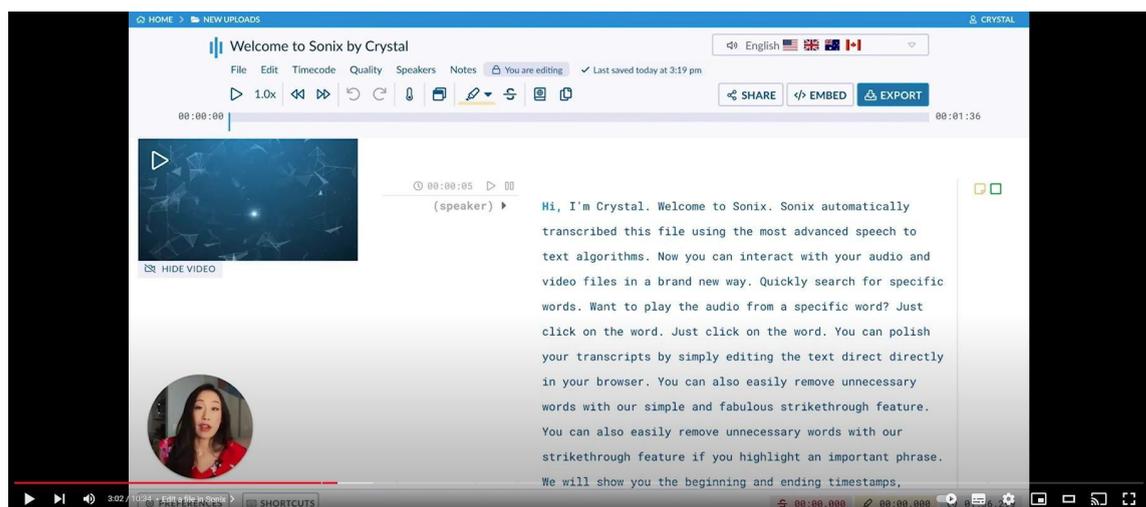


Figura 16 — Vídeo de demonstração de como utilizar a ferramenta Sonix. Fonte: YouTube. Disponível em: https://www.youtube.com/watch?v=_1UXqtclh30

5.4.9 Audext

Com um design mais simples e intuitivo, a Audext aposta em um produto eficiente e barato para a disputa de mercado. Analisando os indicadores, ele atinge o nível mediano neles todos. Não tem a melhor acurácia (80%), mas é uma porcentagem considerada boa, assim como um período de teste razoável com trinta minutos de áudio transcrito, um excelente suporte para idiomas (60 idiomas) e um ótimo preço com R\$0,41 por minuto transcrito. Considerando isso, podemos dizer que é a ferramenta mais ponderada entre os softwares

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

listados.

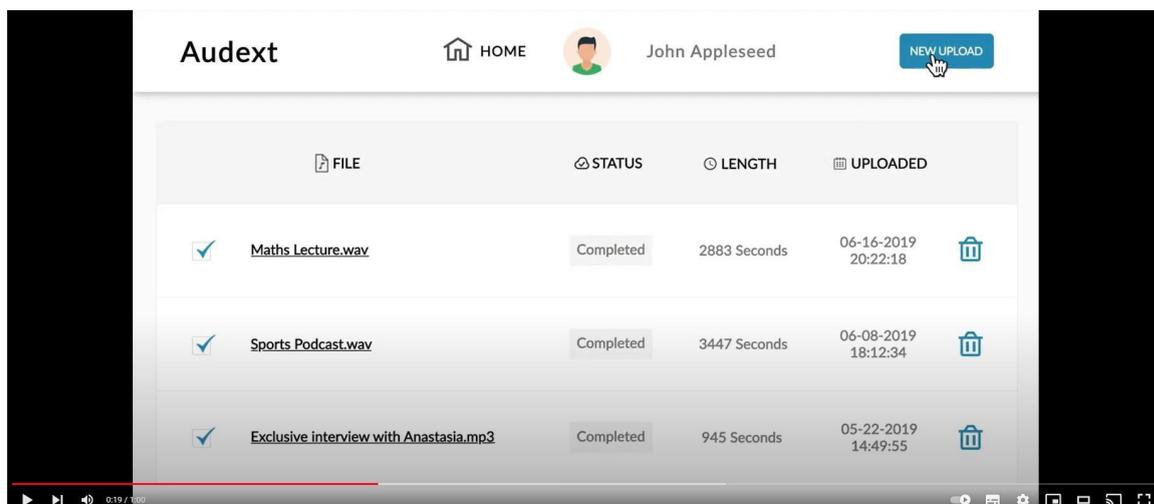


Figura 17 — Vídeo de demonstração de como utilizar a ferramenta Audext. Fonte: YouTube. Disponível em: https://www.youtube.com/watch?v=_1UXqtclh30

5.4.10 Comparação entre softwares de transcrição de áudio em texto

TABELA COMPARATIVA DE SOFTWARES TRANSCRITORES

Software	Transc. em Tempo Real	Suporte p/ Línguas	Teste Grátis	Preço (USD/min)	Acurácia	Comp. do Áudio (min)	Custo do Plano Básico (R\$/Hora)
ISA	Sim	50	Ilimitado	0	80%	Ilimitado	R\$ 0,00
Sonix	Não	30	30min	6	95%	60	R\$ 1.886,40
Trint	Não	31	540min	0,05	87%	180	R\$ 15,72
Audext	Não	60	30min	0,08	80%	60	R\$ 25,15
Amberscript	Não	39	10min	10	85%	60	R\$ 3.144,00
Transcribe	Não	86	1min	0,1	87%	60	R\$ 31,44
Descript	Não	1	180min	2	95%	180	R\$ 628,80
Happy Scribe	Não	120	10min	0,23	85%	5	R\$ 72,31
Rev	Não	1	45min	0,25	80%	5	R\$ 78,60
Temi	Não	1	45min	0,25	92%	Ilimitado	R\$ 78,60

Tabela 2 – Tabela comparativa entre o ISA e os outros softwares transcritores de texto.

Uma Biblioteca do Python é uma coleção de módulos de script acessíveis a um programa Python para simplificar o processo de programação e remover a necessidade de reescrever os comandos mais usados. Eles podem ser usados chamando-os / importando-os no início de um script.

Em comparação com os outros softwares de transcrição de áudio em texto disponíveis no mercado o ISA possui ampla vantagem, o maior destaque é ser o único software que faz transcrição em tempo real e com isso permite um leque maior de utilizações e usuários, como, por exemplo, seu uso em vídeo chamadas por deficientes auditivos, legendagem de produtos audiovisuais por produtores independentes e criadores de conteúdo online e ainda utilizada por TVs comunitárias e de ONGs sendo possível a integração com os softwares de transmissão.

Outro destaque desta é seu uso ilimitado e gratuito, ponto que demonstra sua superioridade perante as outras ferramentas de mercado, possibilitando assim o uso por toda e qualquer pessoa. Para comparação, utilizando um cenário, de uma TV Comunitária que precisa de ferramenta de legendagem em pelo menos vinte horas de sua programação, e hipoteticamente a Trint, que é a ferramenta mais barata, possuir o recurso de transcrição em tempo real, a TV gastaria algo em torno de R\$314,40, por dia, o que daria cerca de R\$9.432 mensais. Algo totalmente inviável para uma TV comunitária. Utilizando o ISA essa TV poderia fazer o uso da ferramenta o quanto for necessário, gratuitamente.

Sua acurácia não é das melhores, mas se iguala aos que os outros softwares oferecem. O mesmo para a gama de idiomas de transcrição, pois utilizando o motor de tradução do Google, o ISA tem disponibilidade para transcrever para mais de cinquenta idiomas, o deixando em uma posição privilegiada perante os demais.

Em uma classificação geral, o ISA ficaria em primeiro lugar no custo, comprimento de áudio, teste grátis e transcrição em tempo real. Estaria empatado em sexto lugar em acurácia e em quarto para suporte de idiomas. O que o colocaria em primeiro lugar levando em consideração os quesitos usabilidade e custo-benefício. Os valores em dólares foram convertidos em uma razão de 5,24 reais brasileiros para cada dólar americano. Cotação do dia 9 de fevereiro de 2022.

5.5 ISA — INTERFACE DE SUPORTE À ACESSIBILIDADE

Com o intuito de auxiliar o principalmente o público com alguma categoria de deficiência auditiva, foi idealizado o desenvolvimento de uma interface de auxílio digital para transcrever todo e qualquer som proveniente do computador pessoal de quem o utiliza. Entretanto, para contextualizar, devemos iniciar descrevendo o que é uma interface. Sua etimologia é composta pelo prefixo em latim *inter*, que significa estar entre ou no meio de, e o substantivo comum *face*, sendo a superfície ou face de algo, sendo que a junção desses termos [*interface*] é aquilo que está entre duas ou mais faces e/ou superfícies.

Transmutando para a linguagem a característica do meio digital, vê-se que é “o elemento que proporciona uma ligação física ou lógica entre dois sistemas, ou partes de um sistema que não poderiam estar conectados diretamente” (SAWAYA, 1999, p. 239). Esses sistemas podem ser de vários tipos, como: gráfica, de áudio, de voz, de texto, entre outras. Porém, todos eles têm em comum um ponto crucial e fundamental para essa ligação, o usuário.

Um dos objetivos desse projeto é construir uma interface híbrida, a qual possui características de várias outras interfaces tais quais: gráfica, por meio da interação com o usuário, a de áudio, na captação dos sons, de texto, na transcrição dos sons e até a de voz a ser utilizada principalmente em suporte a videochamadas com pessoas que não saibam LIBRAS.

O modo de utilização dessa interface será o mais simples e intuitivo possível, no qual o usuário irá executar o programa e a partir daí ele irá começar a transcrever qualquer tipo de som fonético proveniente do seu computador, seja uma música, vídeo ou som que possua fala ou vozes. Essa transcrição será efetuada simultaneamente pelo programa apoiado pela biblioteca de códigos *Speech Recognition* em segundo plano que a exibirá em uma janela de interface gráfica na tela do computador.

A ISA foi desenvolvida e testada para ter compatibilidade com a maioria dos computadores, salvaguarda aqueles que possuem módulos tecnológicos, como os *drives*⁴ de execução de áudio, por exemplo, incompatíveis com a mesma. E sua distribuição será gratuitamente por um link aberto do

repositório pessoal do autor desse projeto, portanto, aqueles que desejarem utilizar, deverão fazer o download do protótipo de código aberto e executá-lo em seu computador.

5.5.1 ISA versus LSE/Closed Caption

Atualmente o mais próximo que temos de uma interface de auxílio às pessoas surdas é a LSE, ou legendagem para surdos e ensurdecidos. Contudo, esse tipo de auxílio é apenas utilizado em alguns tipos de produtos audiovisuais como programas de televisão, filmes, séries e minisséries, todos eles reproduzidos em TV aberta ou fechada, orientada e exigida pela Lei Brasileira de Inclusão da Pessoa com Deficiência n.º 13.146 de 6 de julho de 2015 (BRASIL, 2015).

Todavia, essa ferramenta ainda é limitada apenas a esses produtos, sendo que o usuário fica à mercê de quem está reproduzindo esse conteúdo para inseri-la. Posto isto, é impossível a adaptação tecnológica dessa ferramenta para o auxílio em outro tipo de produto audiovisual ou até mesmo para a programação televisiva, pois necessita de hardwares específicos, sendo eles de custo de aquisição alto, além do custo de implementação e manutenção dele, sem contar na necessidade de profissionais qualificados para essas etapas.

Essa adaptabilidade vai além da assistência, ela é necessária e essencial para os dias de hoje. Tendo em vista que, segundo a Agência Brasil, em um censo realizado em 2020, 81% da população brasileira com mais de dez anos têm internet em casa e cerca de 86% possuem computadores e incluída nesses números, com certeza temos grande parte e, por que não, toda a população surda do Brasil. Por esse motivo, uma ferramenta de auxílio compatível com esses dispositivos é de gigantesca valia, visto que em sua maioria, os outros softwares não possuem tradução para português e muitos deles sendo de difícil implementação e utilização, e que nesse contexto, vai muito além do auxílio, sendo encarada como uma ferramenta de inclusão que se olharmos de forma analítica, podemos considerar que o desenvolvimento de uma, está há alguns anos, atrasado aqui no Brasil. A ISA tem por principal objetivo ser essa ferramenta, essa interface, que ligará o público surdo à

acessibilidade em produtos audiovisuais de livre escolha, ou no caso, da escolha do usuário, sendo essa sua maior vantagem.

Desse modo, o usuário poderá consumir o conteúdo audiovisual de seu desejo e ter a assistência acessível a todo instante, sem depender de terceiros, apenas do anseio de utilizar a ISA.

5.5.2 Benefícios Gerais da ISA

De modo geral, além de sua funcionalidade, a ISA possui outros benefícios, sendo o principal deles a compatibilidade com a maioria dos computadores que são utilizados hoje, além de ser *Open Source*, que:

[] significa “código aberto” e se refere ao código-fonte de um site ou aplicativo. Nesse sentido, sua linguagem de programação pode ser vista por qualquer um, que pode adaptá-la para objetivos variados. Na maioria das vezes, essas ferramentas são compartilhadas na rede pelos desenvolvedores, permitindo acesso sem restrições a qualquer pessoa.” (BLOG TOTVS, 2018).

Outro benefício não menos importante é a sua distribuição gratuita, sendo fator crucial para a inclusão, fazendo com que toda e qualquer pessoa que desejar utilizar a ferramenta, tenha acesso a ela através de um repositório intuitivo para o download. Também é uma funcionalidade da ISA que pode ser vista como vantagem, o usuário poder desligar a assistente, fechando-o a qualquer momento. Essa funcionalidade visa o público que possui deficiência auditiva de nível moderada e leve, que ainda conseguem ouvir algumas faixas de som (medidas em decibéis e Hertz) e queiram ouvir o som, normalmente uma música, sem a assistência da ferramenta, semelhante ao *Closed Caption*, salvaguarda quando o mesmo está fixado permanentemente no produto audiovisual pois foi inserido na produção do mesmo.

6 CÓDIGO FONTE

Iniciamos o desenvolvimento da programação do protótipo, importando a principal biblioteca, a *Speech Recognition*, que irá reger todo o desenvolvimento do arquétipo.

```

1 ##### BIBLIOTECAS #####
2
3 import speech_recognition as sr
4

```

Figura 18 — Linha de código de importação da biblioteca de transcrição.

Após isso, precisamos dar função ao módulo de apoio, o qual reconhece o driver do microfone e do alto-falante do computador do usuário, para fazer a captura dos espectros dos áudios para um ambiente virtual (sem utilizar o computador do usuário), um tratamento do áudio, o acionamento da IA para o reconhecimento de fonemas e a transcrição deles, tudo isso em milésimos de segundos aproximando-se muito do “tempo real”.

```

5 ##### MÓDULO DE APOIO #####
6
7 def reconheceraudiomicrofone(reconhecedor, microfone):
8
9     with microfone as somprincipal:
10         reconhecedor.adjust_for_ambient_noise(somprincipal)
11
12         audio = reconhecedor.listen(somprincipal)
13
14         response = {
15             "sucesso": True,
16             "erro": None,
17             "transcricao": None
18         }
19
20         try:
21             response["transcricao"] = reconhecedor.recognize_google(audio, language='pt_BR')
22
23         except sr.RequestError:
24             response["sucesso"] = False
25             response["erro"] = "# Programa Indisponível"
26
27         except sr.UnknownValueError:
28             response["erro"] = "# Áudio Não Reconhecido"
29
30         return response

```

Figura 19 — Códigos do módulo de apoio que fazem o background da transcrição

Por fim, precisamos itinerar, transitar, trafegar e acompanhar os comandos para que o protótipo efetue a transcrição e traga essa mesma transcrição para fora do software através de uma interface, exibindo assim as palavras que foram transcritas do áudio captado via os drivers supracitados. Acrescentamos também algumas orientações de início de utilização e para caso haja algum erro na execução, como podemos ver nas figuras 20, 21 e 22, abaixo:

```

33 ##### MÓDULO PRINCIPAL DE RECONHECIMENTO DE ÁUDIO #####
34
35 if __name__ == "__main__":
36
37     reconhecedor = sr.Recognizer()
38     microfone = sr.Microphone()
39
40     NFALA = 2
41     NSEMFALA = 2
42
43     print('\n ### ISA - Interface de Suporte a Acessibilidade - PROGRAMA FUNCIONANDO ###')
44
45     for i in range(NFALA):
46
47         for j in range(NSEMFALA):
48
49             print('\nAguarde um momento estou identificando o áudio...\n'.format(i+1))
50
51             somteste = reconheceaudiomicrofone(reconhecedor, microfone)
52
53             if somteste["transcricao"]:
54                 break
55
56             if not somteste["sucesso"]:
57                 print("# Áudio não identificado, tentando novamente...")
58                 break
59
60             if somteste["erro"]:
61                 print("ERRO: {}".format(somteste["erro"]))
62                 break
63
64             print("# Áudio Identificado: {}".format(somteste["transcricao"]))

```

Figura 20 — Códigos que efetuam de fato a transcrição do áudio em texto.

```

### ISA - Interface de Suporte a Acessibilidade - PROGRAMA FUNCIONANDO ###

Aguarde um momento estou identificando o áudio...

# Áudio Identificado: este é um teste utilizando o protótipo Isa para reconhecimento de fala via microfone

```

Figura 21 — Protótipo ISA em funcionamento, teste de transcrição via Microfone.

```

### ISA - Interface de Suporte a Acessibilidade - PROGRAMA FUNCIONANDO ###

Aguarde um momento estou identificando o áudio...

ERRO: # Áudio Não Reconhecido

# Áudio Identificado: None

Aguarde um momento estou identificando o áudio...

# Áudio Identificado: primo Tribunal Federal Alexandre de Moraes revogou neste domingo a ordem de bloqueio o aplicativo telegram a revogação ocorre após o cumprimento de determinações da corte que estavam pendentes a decisão de bloquear a plataforma em todo o país publicada na última sexta-feira não chegou a ser implementada e o aplicativo seguiu funcionando normalmente no fim de semana após o anúncio da suspensão o fundador do telegram pediu desculpas ao STF alegando que as relações se deviam a uma falha de comunicação por e-mail as extraviados e pediu mais tempo para ficar em dia com a solicitação

```

Figura 22 — Protótipo ISA em funcionamento, teste de transcrição a partir de vídeo do YouTube via alto-falante.

7 CONCLUSÃO

Em sua versão desenvolvida e liberada atualmente, a versão Beta 0.7.6, o protótipo ainda necessita de várias melhorias técnicas, como a assertividade da fala em relação à pontuação, a identificação de personagens e tipos de som para que na exibição esteja determinado quem é o interlocutor no momento em que o som é emitido.

Outra melhoria é na compatibilidade com outros sistemas operacionais que não o Windows. Por mais que seja o sistema operacional mais popular atualmente (NOLETO [s.d.]), o intuito é que seja uma ferramenta acessível a

todos os públicos, e para isso os arranjos de compatibilidade devem ser feitos para que o protótipo funcione em todo e qualquer tipo de computador pessoal. O potencial de desenvolvimento do produto é bastante grande, visto que a taxa de acurácia do reconhecimento ainda está em 80%, podendo ser melhorada até próximo dos 100% com a adição de módulos de Machine Learning e Deep Learning. Isso melhoraria a naturalidade do reconhecimento do som, trazendo uma transcrição mais fluida e assertiva com relação à gramática. Outro ponto de melhoria é na questão visual, visando facilitar a visibilidade dos textos, trabalhar em tons de alto contraste para pessoas com baixa visão, atrelada a deficiência auditiva.

Dado que em seu estado atual o protótipo é utilizável por qualquer usuário ou produtor audiovisual que tenha um computador com sistema operacional com Windows, bastando apenas ser efetuado o download do mesmo em repositório de código aberto, disponibilizado na seção 1.1 deste mesmo documento, e cumprindo com as funções projetadas, podemos afirmar que o presente estudo e desenvolvimento obteve resultado satisfatório.

Espero que em breve possamos ver o protótipo já em sua versão final, utilizado para o fomento do movimento de acessibilidade no Brasil.

8 REFERÊNCIAS

NAVES, S. B. et al. **Guia para produções audiovisuais acessíveis**. Brasília: Secretaria do Audiovisual/ Ministério da Cultura, 2016.

ARAÚJO, V. L. S. **Por um modelo de legendagem para surdos no Brasil**. Tradução & Comunicação, n. 17, p. 59-76, 2008.

CAMINOS, A. MÉDOLA, A. S.; SUING, A. (orgs.). **A nova televisão – do Youtube ao Netflix**. Aveiro: Ria Editorial, 2019.

Convenção sobre os Direitos das Pessoas com Deficiência (2007). Convenção sobre os direitos das pessoas com deficiência. Vitória: Ministério Público do Trabalho, Projeto PCD Legal, 2014.

BARIFFI, F. J. et al. **La accesibilidad universal en los medios audiovisuales de comunicación**. Madrid: Real Patronato Sobre Discapacidad; CESyA, 2008.

MACHADO, F. O. **Acessibilidade na televisão digital: estudo para uma política de audiodescrição na televisão brasileira**. 2011. 180 f. Dissertação

(mestrado) - Universidade Estadual Paulista, Faculdade de Arquitetura, Artes e Comunicação, 2011.

FRYER, L.: **An introduction to audio description: a practical guide.** Oxford: Routledge, 2016

MOTTA, L. M. V. de M. M.; ROMEU FILHO, P. (Org.). **Audiodescrição: transformando imagens em palavras.** São Paulo: Secretaria de Estado dos Direitos da Pessoa com Deficiência, 2010.

DELGADO, F. U. **Accesibilidad a la TDT en España para personas con discapacidad sensorial.** (Tesis de doctorado). Madrid: Universidad Carlos III de Madrid, 2008.

MATAMALA, A.; ORERO, P. (eds) **Researching audio description new approaches.** London: Palgrave Macmillan, 2016.

São Paulo (Cidade). Secretaria Municipal da Pessoa com Deficiência - SMPED. **Conhecer para Incluir a Pessoa com Deficiência.** São Paulo: SMPED; Editora Mais Diferenças, 2020.

SAWAYA, Marcia Regina. **Dicionário de informática e internet.** São Paulo: Nobel, 1999.

OPEN Source (Código Aberto): Como funciona e suas vantagens. Blog Totvs, 2021. Disponível em: <<https://www.totvs.com/blog/developers/como-funciona-um-software-open-source/>>. Acesso em: 04 de out. de 2021.

PÉREZ-GONZALEZ, Luis. **Audiovisual Translation Theories, Methods And Issues.** New York: Routledge, 2014.

BARBOZA, L. **Inteligência Artificial Com R E Python Para Iniciantes.** S.L.: Novas Edições Acadêmicas, 2020.

The Invisible Subtiter - A Documentary (SDH Subtitles included). Disponível em: <<https://www.youtube.com/watch?v=Pz75i6EsOto>>. Acesso em: 5 jan. 2022.

FULLERTON, J. Introduction: Local Film. **Film History: An International Journal**, 2005.Vol.23. Edição 1-2. Págs. 81-94.

SLIDE, A. **The new historical dictionary of the American film industry.** Lanham, Md. London: Scarecrow, 2001.

BARNES, J. **Victorian Film History. The beginnings of cinema in England.** London: David & Charles. 1976. Disponível em: <<http://www.victorian-cinema.net/paul>> Acesso em: 5 jan. 2022.

The British Film Catalogue. 1895-1970. Londres. Davi & Carlos. ISBN 0-715305572-4. Disponível em: <<http://www.britishpictures.com/books/catalogue.htm>> Acesso em: 5 jan. 2022.

HEPWORTH, C. M. **Came the dawn : Memories of a film pioneer**. London: Phoenix, 1951. Disponível em: <<http://www.screenonline.org.uk/people/id/450004/>> Acesso em: 5 jan. 2022.

THE BENSHI TRADITION: CINEMA=PERFORMANCE. Disponível em: <<http://www.altx.com/interzones/kino2/benshi.html>>. Acesso em: 12 jan. 2022.

FONG, G. C. F.; AU, K. K. L. **Dubbing and subtitling in a world context**. Hong Kong: Chinese University Press, 2009.

HENRIK GOTTLIEB. **Screen translation: eight studies in subtitling, dubbing and voice-over**. Copenhagen: Henrik Gottlieb/University Of Copenhagen, 2005

A Wee corner of Showbiz. The Emporia Gazette. Emporia. Kansas. 1960. p 2.

WEINBERG, H. G. **A Manhattan odyssey : a memoir**. New York: Anthology Film Archives, 1982. p108.

FREEDMAN, S. G. HERMAN G. WEINBERG, WRITER AND FOREIGN FILM TRANSLATOR. **The New York Times**, 8 nov. 1983.

Ways to Strength and Beauty. Disponível em: <https://en.wikipedia.org/wiki/Ways_to_Strength_and_Beauty>. Acesso em: 12 jan. 2022.

CROSLAND, A. et al. **The Jazz Singer**. Disponível em: <https://www.imdb.com/title/tt0018037/?ref_=nv_sr_2>. Acesso em: 12 jan. 2022.

N. Kaufmann & W. Prager: **Wege zu Kraft und Schönheit (1925)**. Disponível em: <<https://www.youtube.com/watch?v=RsvR90zBV-w>>. Acesso em: 12 jan. 2022.

RUGGLES, W. et al. **The Plastic Age**. Disponível em: <https://www.imdb.com/title/tt0016226/?ref_=nv_sr_1>. Acesso em: 13 jan. 2022.

BACON, L. et al. **The Singing Fool**. Disponível em: <<https://www.imdb.com/title/tt0019388/>>. Acesso em: 13 jan. 2022.

NORNES, M. **Cinema babel : translating global cinema**. p 21. Minneapolis; London: University Of Minnesota Press, Cop, 2007.

STERNBERG, J. VON et al. **Morocco**. Disponível em: <<https://www.imdb.com/title/tt0021156/>>. Acesso em: 13 jan. 2022.

A Record De Shanghai (1938). Disponível em: <<https://www.youtube.com/watch?v=sZVkusbuH-o>>. Acesso em: 13 jan. 2022.

IVARSSON, J. **A Short Technical History of Subtitles in Europe**. Transedit,

2004.

The Rank Organization. Disponível em: https://en.wikipedia.org/wiki/Rank_Organisation>. Acesso em: 13 jan. 2022.

BIO | Maï Boiron Sous-titres et doublage. Disponível em: <https://www.maiboiron.fr/bio>>. Acesso em: 14 jan. 2022.

AUBOYER, D.; CATELLANI, C. **Mecanisme pour la gravure par un faisceau laser de sous-titres simultanément sur une ou plusieurs images d'un film cinematographique.** Disponível em: <https://patents.google.com/patent/FR2855884B1/en?inventor=Denis+Auboyer>>. Acesso em: 14 jan. 2022.

EDCF | Guide to Digital Cinema Mastering. Disponível em: <https://edcf.net/articles-publications/>>. Acesso em: 14 jan. 2022.

Illustrated Guide. Disponível em: <https://registry-page.isdcf.com/illustratedguide/>>. Acesso em: 14 jan. 2022.
NORNES, M. **Cinema babel: translating global cinema.** p 216-217.

Minneapolis ; London: University Of Minnesota Press, Cop, 2007.

Instituto Europeu de Patentes. Requerente: Thomson Licensing 92130 Issy-Les- Moulieaux (FR). Pedido de Patente EP 2 938 087 A1. pp.7-8.

Sir Howard Stringer, Sony Chairman and CEO, Presents "The CES Seven," Imperatives for Creating the Ultimate User Experience, at 2009 Consumer Electronics Show. Disponível em: https://www.sony.com/content/sony/en/en_us/SCA/company-news/press-releases/sony-electronics/2009/sir-howard-stringer-sony-chairman-and-ceo-presents-the-ces-seven-imperatives-for-creating-the-ultimate-user-experience-at-2009-consumer-electronics-show.html>. Acesso em: 18 jan. 2022.

Invisible Subtitles. Disponível em: <https://www.3dexperience.co.uk/subtitles.html>>. Acesso em: 18 jan. 2022.

User's Guide. [s.l.: s.n.]. Disponível em: <https://files.support.epson.com/docid/cpd4/cpd40542.pdf>>. Acesso em: 18 jan. 2022.

Off-Screen Cinema Subtitling System. Disponível em: <https://www.youtube.com/watch?v=b6uDWI4wVFg>>. Acesso em: 18 jan. 2022.

NORNES, M. **Cinema babel : translating global cinema.** p 25. Minneapolis ; London: University Of Minnesota Press, Cop, 2007.

Flycz Blog, **intertitles,** 2014, Disponível em: <https://flyczba.wordpress.com/category/intertitles/>>. Acesso em: 12 jan. 2022.

TIOBE Index | TIOBE - The Software Quality Company. Disponível em: <https://www.tiobe.com/tiobe-index/>>. Acesso em: 18 jan. 2022.

Instituto Locomotiva, **Raio-X da Surdez no Brasil**, Jornal O Estado de São Paulo, 2019. Disponível em: <https://brasil.estadao.com.br/blogs/vencer-limites/wp-content/uploads/sites/189/2019/09/raioXSurdez_InstitutoLocomotiva_30setembro2019_blogVencerLimites.pdf>. Acesso em: 10 mai. 2021.

O que é Inteligência artificial? Como funciona, exemplos e aplicações. Disponível em: <<https://www.totvs.com/blog/inovacoes/o-que-e-inteligencia-artificial/>>.

History and Philosophy of Python | Python Tutorial | python-course.eu. Disponível em: <<https://python-course.eu/python-tutorial/history-and-philosophy-of-python.php>>. Acesso em: 12 jan. 2022.

What is a Stenographer and How Does Stenography Work? | Naegeli. Disponível em: <<https://www.naegeliusa.com/blog/what-is-a-stenographer>>. Acesso em: 13 jan. 2022.

Court Reporters : Occupational Outlook Handbook: : U.S. Bureau of Labor Statistics. Disponível em: <<https://www.bls.gov/ooh/legal/court-reporters.htm>>. Acesso em: 13 jan. 2022.

IBGE Censo 2010. Disponível em: <<https://censo2010.ibge.gov.br/sinopse/index.php?dados=P6&uf=00>>. Acesso em: 12 jan. 2022.

ANATEL. Anatel - Portaria nº 310, de 27 de junho de 2006. Disponível em: <<https://informacoes.anatel.gov.br/legislacao/normas-do-mc/442-portaria-310>>. Acesso em: 12 jan. 2022.

Closed caption: saiba como funcionam as legendas automáticas nas TVs. Disponível em: <<https://www.uol.com.br/tilt/noticias/redacao/2012/04/26/closed-caption-saiba-como-funcionam-as-legendas-automaticas-nas-tvs.htm>>. Acesso em: 18 jan. 2022.

O que são drivers? Entenda o que são e como eles funcionam nos PCs . Disponível em: <<https://canaltech.com.br/produtos/drivers-entenda-o-que-sao-e-como-eles-funcionam-195604/#:~:text=De%20forma%20simples%2C%20o%20driver>>. Acesso em: 18 jan. 2022.

BRASIL. Decreto n.º 13.146 de 6 de julho de 2015. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). **Diário Oficial da União**: Seção 1, Brasília, 05 de julho de 2015, p. 2 Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm>. Acesso em: 22 jun. 2021.

Kinetoscope. Disponível em: <<https://en.wikipedia.org/wiki/Kinetoscope#Kinetophone>>. Acesso em: 15 jan. 2022.

NEILSON, K.; MANVELL, R.; WOLLENBERG, H. H. **The Penguin film review.** Harmondsworth: Penguin, 1948.

Subtitling for Cinema: A Brief History | The Artifice. Disponível em: <<https://the-artifice.com/subtitling-cinema-history/>>. Acesso em: 12 jan. 2022.

[HTTPS://MUBI.COM/CAST/CECIL-M-HEPWORTH](https://mubi.com/cast/cecil-m-hepworth); **How It Feels to Be Run Over.** Disponível em: <<https://mubi.com/pt/films/how-it-feels-to-be-run-over>>. Acesso em: 12 jan. 2022.

ETTLEMAN, T. **The First Film to Use Intertitles Was Also the First Dickens Adaptation.** Disponível em: <<https://trettleman.medium.com/the-first-film-to-use-intertitles-was-also-the-first-dickens-adaptation-cfd906684011>>. Acesso em: 12 jan. 2022.

Alice in Wonderland (Film, Fantasy): Reviews, Ratings, Cast and Crew – Rate Your Music. Disponível em: <https://rateyourmusic.com/film/alice_in_wonderland_f2/>. Acesso em: 12 jan. 2022.

The first movie with verbal texting - “College Chums” 1907 by Edwin S Porter Thomas Edison. Disponível em: <<https://www.youtube.com/watch?v=5JS4rLQmtA>>. Acesso em: 12 jan. 2022.

Silent Era: Progressive Silent Film List. Disponível em: <<http://www.silentera.com/PSFL/data/C/ChamberMystery1920.html>>. Acesso em: 12 jan. 2022.

HEBB, D. **The Organization of Behavior A NEUROPSYCHOLOGICAL THEORY.** [s.l: s.n.], 1949. Disponível em: <https://pure.mpg.de/rest/items/item_2346268_3/component/file_2346267/content>. Acesso em: 15 jan. 2022.

11.2 Samuel’s Checkers Player. Disponível em: <<http://www.incompleteideas.net/book/ebook/node109.html>>. Acesso em: 15 jan. 2022.

Step by Step: Alpha Beta Pruning. Disponível em: <<https://www.youtube.com/watch?v=xBXHz4Gbdo>>. Acesso em: 15 jan. 2022.

HINTZE, A. **Computers to humans: Shall we play a game?** Disponível em: <<https://theconversation.com/computers-to-humans-shall-we-play-a-game-77383>>. Acesso em: 15 jan. 2022.

NOLETO, C. **O que é Sistema Operacional? O guia completo!** Disponível em: <<https://blog.betrybe.com/tecnologia/sistema-operacional-tudo-sobre/>>. Acesso em: 15 jan. 2022.

JUSTIN, B. **Introduction to Perceptron: Neural Network**. Disponível em: <<https://blog.knoldus.com/introduction-to-perceptron-neural-network/>>. Acesso em: 16 jan. 2022.

COVER, T. M., HART P. E. **This Week's Citation Classic**. 1967. Disponível em: <<http://garfield.library.upenn.edu/classics1982/A1982NF37700001.pdf>>. Acesso em: 16 jan. 2022.

NIELSEN, M. A. **Neural Networks and Deep Learning**. Disponível em: <<http://neuralnetworksanddeeplearning.com/chap2.html>>. Acesso em: 16 jan. 2022.

KURENKOV, R. **A Brief History of Neural Nets and Deep Learning**. Disponível em: <<https://www.skynettoday.com/overviews/neural-net-history>>. Acesso em: 16 jan. 2022.

SCHAPIRE, R. E. **The strength of weak learnability**. 30th Annual Symposium on Foundations of Computer Science, 1989.

BROWNLEE, J. **Boosting and AdaBoost for Machine Learning**, 2016. Disponível em: <<https://machinelearningmastery.com/boosting-and-adaboost-for-machine-learning/>>. Acesso em: 16 jan. 2022.

MAYO, B. **Apple details how the HomePod uses machine learning to ignore noisy surroundings for Siri requests - 9to5Mac**, 2018. Disponível em: <<https://9to5mac.com/2018/12/03/homepod-machine-learning-siri-speech-recognition/>>. Acesso em: 16 jan. 2022.

Face Recognition Grand Challenge (FRGC). Disponível em: <<https://www.nist.gov/programs-projects/face-recognition-grand-challenge-frgc>>. Acesso em: 16 jan. 2022.

MONTTI, R. **FLAN: Google Research Develops Better Machine Learning**. Disponível em: <<https://www.searchenginejournal.com/flan-google-research-finds-better-machine-learning/428732/#close>>. Acesso em: 16 jan. 2022.

EDDY, S. R. **What is a hidden Markov model?** Nature Biotechnology, v. 22, n. 10, p. 1315–1316, out. 2004.

MOSKVITCH, K. **The machines that learned to listen**. Disponível em: <<https://www.bbc.com/future/article/20170214-the-machines-that-learned-to-listen>>. Acesso em: 18 jan. 2022.

HowTo Transcribe Video to Text (2021). Disponível em: <https://www.youtube.com/watch?v=_1UXqtclh30>. Acesso em: 18 jan. 2022.

Introducing Premium Service Captions | Rev Pro. Disponível em: <<https://www.youtube.com/watch?v=dwFZgz4eDKg>>. Acesso em: 18 jan. 2022.

How to order Automatic transcription? | Disponível em: <https://www.youtube.com/watch?v=b_PSCzX4hsg>. Acesso em: 18 jan. 2022.

Transcribe.Wreally - Digital Transcription Application Demonstration, Video 1. Disponível em: <<https://www.youtube.com/watch?v=TsUfITTOvu4>>. Acesso em: 18 jan. 2022.

TRINT. **Realtime Transcription by TrintYouTube**, 12 nov. 2019. Disponível em: <<https://www.youtube.com/watch?v=90iWFDxjF-c>>. Acesso em: 18 jan. 2022.

How to Transcribe using Temi.com. Disponível em: <https://www.youtube.com/watch?v=_oJoAicyZD0>. Acesso em: 18 jan. 2022.

Descript Transcription Tutorial | How To Use Descript For Automatic Transcription. Disponível em: <<https://www.youtube.com/watch?v=Ra50ReVMBsw>>. Acesso em: 18 jan. 2022.

Introduction to Sonix: Learn the basics. Disponível em: <<https://www.youtube.com/watch?v=9sPJ9szix8Y>>. Acesso em: 18 jan. 2022.

How to automatically transcribe audio (Convert Audio) to text in 2019? disponível em: <<https://www.youtube.com/watch?v=ssJnvYaag3A>>. Acesso em: 18 jan. 2022.

11.2 Samuel's Checkers Player. Disponível em: <<http://www.incompleteideas.net/book/ebook/node109.html>>. Acesso em: 15 jan. 2022.

A Record De Shanghai (1938). Disponível em: <<https://www.youtube.com/watch?v=sZVkusbuH-o>>. Acesso em: 13 jan. 2022.

A Wee corner of Showbiz. The Emporia Gazette. Emporia. Kansas. 1960. p 2.

Alice in Wonderland (Film, Fantasy): Reviews, Ratings, Cast and Crew – Rate Your Music. Disponível em: <https://rateyourmusic.com/film/alice_in_wonderland_f2/>. Acesso em: 12 jan. 2022.

ANATEL. Anatel - Portaria no 310, de 27 de junho de 2006. Disponível em: <<https://informacoes.anatel.gov.br/legislacao/normas-do-mc/442-portaria-310>>. Acesso em: 12 jan. 2022.

ARAÚJO, V. L. S. Por um modelo de legendagem para surdos no Brasil. Tradução & Comunicação, n. 17, p. 59-76, 2008.

AUBOYER, D.; CATELLANI, C. Mecanisme pour la gravure par un faisceau laser de sous-titres simultanément sur une ou plusieurs images d'un film cinematographique. Disponível em: <<https://patents.google.com/patent/FR2855884B1/en?inventor=Denis+Auboyer>>. Acesso em: 14 jan. 2022.

BACON, L. et al. The Singing Fool. Disponível em:

<<https://www.imdb.com/title/tt0019388/>>. Acesso em: 13 jan. 2022.

BARBOZA, L. Inteligência Artificial Com R E Python Para Iniciantes. S.L.: Novas Edições Acadêmicas, 2020.

BARIFFI, F. J. et al. La accesibilidad universal en los medios audiovisuales de comunicación. Madrid: Real Patronato Sobre Discapacidad; CESyA, 2008.

BARNES, J. Victorian Film History. The beginnings of cinema in England. London: David & Charles. 1976. Disponível em: <<http://www.victorian-cinema.net/paul>> Acesso em: 5 jan. 2022.

BIO | Maï Boiron Sous-titres et doublage. Disponível em: <<https://www.maiboiron.fr/bio>>. Acesso em: 14 jan. 2022.

BRASIL. Decreto n.º 13.146 de 6 de julho de 2015. Institui a Lei Brasileira de Inclusão da Pessoa com Deficiência (Estatuto da Pessoa com Deficiência). Diário Oficial da União: Seção 1, Brasília, 05 de julho de 2015, p. 2 Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2015/lei/l13146.htm>. Acesso em: 22 jun. 2021.

BROWNLEE, J. Boosting and AdaBoost for Machine Learning, 2016. Disponível em: <<https://machinelearningmastery.com/boosting-and-adaboost-for-machine-learning/>>. Acesso em: 16 jan. 2022.

CAMINOS, A. MÉDOLA, A. S.; SUING, A. (orgs.). A nova televisão – do Youtube ao Netflix. Aveiro: Ria Editorial, 2019.

Closed caption: saiba como funcionam as legendas automáticas nas TVs. Disponível em: <<https://www.uol.com.br/tilt/noticias/redacao/2012/04/26/closed-caption-saiba-como-funcionam-as-legendas-automaticas-nas-tvs.htm>>. Acesso em: 18 jan. 2022.

Convenção sobre os Direitos das Pessoas com Deficiência (2007). Convenção sobre os direitos das pessoas com deficiência. Vitória: Ministério Público do Trabalho, Projeto PCD Legal, 2014.

Court Reporters : Occupational Outlook Handbook: : U.S. Bureau of Labor Statistics. Disponível em: <<https://www.bls.gov/ooh/legal/court-reporters.htm>>. Acesso em: 13 jan. 2022.

COVER, T. M., HART P. E. This Week's Citation Classic. 1967. Disponível em: <<http://garfield.library.upenn.edu/classics1982/A1982NF37700001.pdf>>. Acesso em: 16 jan. 2022.

CROSLAND, A. et al. The Jazz Singer. Disponível em: <https://www.imdb.com/title/tt0018037/?ref_=nv_sr_2>. Acesso em: 12 jan. 2022.

DELGADO, F. U. Accesibilidad a la TDT en España para personas con discapacidad sensorial. (Tesis de doctorado). Madrid: Universidad Carlos III de Madrid, 2008.

Descript Transcription Tutorial | How To Use Descript For Automatic Transcription. Disponível em: <<https://www.youtube.com/watch?v=Ra50ReVMBsw>>. Acesso em: 18 jan. 2022.

EDCF | Guide to Digital Cinema Mastering. Disponível em: <<https://edcf.net/articles-publications/>>. Acesso em: 14 jan. 2022.

EDDY, S. R. What is a hidden Markov model? Nature Biotechnology, v. 22, n. 10, p. 1315–1316, out. 2004.

ETTLEMAN, T. The First Film to Use Intertitles Was Also the First Dickens Adaptation. Disponível em: <<https://trettleman.medium.com/the-first-film-to-use-intertitles-was-also-the-first-dickens-adaptation-cfd906684011>>. Acesso em: 12 jan. 2022.

Face Recognition Grand Challenge (FRGC). Disponível em: <<https://www.nist.gov/programs-projects/face-recognition-grand-challenge-frgc>>. Acesso em: 16 jan. 2022.

Flycz Blog, intertitles, 2014, Disponível em: <<https://flyczba.wordpress.com/category/intertitles/>>. Acesso em: 12 jan. 2022.

FONG, G. C. F.; AU, K. K. L. Dubbing and subtitling in a world context. Hong Kong: Chinese University Press, 2009.

FREEDMAN, S. G. HERMAN G. WEINBERG, WRITER AND FOREIGN FILM

FRYER, L.: An introduction to audio description: a practical guide. Oxford: Routledge, 2016

FULLERTON, J. Introduction: Local Film. Film History: An International Journal, 2005.Vol.23. Edição 1-2. Págs. 81-94.

HEBB, D. The Organization of Behavior A NEUROPSYCHOLOGICAL THEORY. [s.l: s.n.],1949. Disponível em: <https://pure.mpg.de/rest/items/item_2346268_3/component/file_2346267/content>. Acesso em: 15 jan. 2022.

HENRIK GOTTLIEB. Screen translation: eight studies in subtitling, dubbing and voice-over. Copenhagen: Henrik Gottlieb/University Of Copenhagen, 2005

HEPWORTH, C. M. Came the dawn : Memories of a film pioneer. London: Phoenix, 1951. Disponível em: <<http://www.screenonline.org.uk/people/id/450004/>> Acesso em: 5 jan. 2022.

HINTZE, A. Computers to humans: Shall we play a game? Disponível em: <<https://theconversation.com/computers-to-humans-shall-we-play-a-game-77383>>. Acesso em: 15 jan. 2022.

History and Philosophy of Python | Python Tutorial | python-course.eu. Disponível em: <<https://python-course.eu/python-tutorial/history-and-philosophy-of-python.php>>. Acesso em: 12 jan. 2022.

How to automatically transcribe audio (Convert Audio) to text in 2019? disponível em: <<https://www.youtube.com/watch?v=ssJnvYaag3A>>. Acesso em: 18 jan. 2022.

How to order Automatic transcription? | Disponível em: <https://www.youtube.com/watch?v=b_PSCzX4hsg>. Acesso em: 18 jan. 2022.

How to Transcribe using Temi.com. Disponível em: <https://www.youtube.com/watch?v=_oJoAicyZD0>. Acesso em: 18 jan. 2022.

How To Transcribe Video to Text (2021). Disponível em: <https://www.youtube.com/watch?v=_1UXqtclh30>. Acesso em: 18 jan. 2022.

[HTTPS://MUBI.COM/CAST/CECIL-M-HEPWORTH](https://mubi.com/cast/cecil-m-hepworth); How It Feels to Be Run Over. Disponível em: <<https://mubi.com/pt/films/how-it-feels-to-be-run-over>>. Acesso em: 12 jan. 2022.

IBGE Censo 2010. Disponível em: <<https://censo2010.ibge.gov.br/sinopse/index.php?dados=P6&uf=00>>. Acesso em: 12 jan. 2022.

Illustrated Guide. Disponível em: <<https://registry-page.isdcf.com/illustratedguide/>>. Acesso em: 14 jan. 2022.
NORNES, M. Cinema babel: translating global cinema. p 216-217.

Instituto Europeu de Patentes. Requerente: Thomson Licensing 92130 Issy-Les- Moulieaux (FR). Pedido de Patente EP 2 938 087 A1. pp.7-8.

Instituto Locomotiva, Raio-X da Surdez no Brasil, Jornal O Estado de São Paulo, 2019. Disponível em: <https://brasil.estadao.com.br/blogs/vencer-limites/wp-content/uploads/sites/189/2019/09/raioXSurdez_InstitutoLocomotiva_30setembro2019_blogVencerLimites.pdf>. Acesso em: 10 mai. 2021.

Introducing Premium Service Captions | Rev Pro. Disponível em: <<https://www.youtube.com/watch?v=dwFZgz4eDKg>>. Acesso em: 18 jan. 2022.

Introduction to Sonix: Learn the basics. Disponível em: <<https://www.youtube.com/watch?v=9sPJ9szix8Y>>. Acesso em: 18 jan. 2022.

Invisible Subtitles. Disponível em: <<https://www.3dexperience.co.uk/subtitles.html>>. Acesso em: 18 jan. 2022.

IVARSSON, J. A Short Technical History of Subtitles in Europe. Transedit, 2004.

JUSTIN, B. Introduction to Perceptron: Neural Network. Disponível em: <<https://blog.knoldus.com/introduction-to-perceptron-neural-network/>>. Acesso em: 16 jan. 2022.

Kinetoscope. Disponível em: <<https://en.wikipedia.org/wiki/Kinetoscope#Kinetophone>>. Acesso em: 15 jan. 2022.

KURENKOV, R. A Brief History of Neural Nets and Deep Learning. Disponível em: <<https://www.skynettoday.com/overviews/neural-net-history>>. Acesso em: 16 jan. 2022.

MACHADO, F. O. Acessibilidade na televisão digital: estudo para uma política de audiodescrição na televisão brasileira. 2011. 180 f. Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Arquitetura, Artes e Comunicação, 2011.

MATAMALA, A.; ORERO, P. (eds) Researching audio description new approaches. London: Palgrave Macmillan, 2016.

MAYO, B. Apple details how the HomePod uses machine learning to ignore noisy surroundings for Siri requests - 9to5Mac, 2018. Disponível em: <<https://9to5mac.com/2018/12/03/homepod-machine-learning-siri-speech-recognition/>>. Acesso em: 16 jan. 2022.

Minneapolis ; London: University Of Minnesota Press, Cop, 2007.

MONTTI, R. FLAN: Google Research Develops Better Machine Learning. Disponível em: <<https://www.searchenginejournal.com/flan-google-research-finds-better-machine-learning/428732/#close>>. Acesso em: 16 jan. 2022.

MOSKVITCH, K. The machines that learned to listen. Disponível em: <<https://www.bbc.com/future/article/20170214-the-machines-that-learned-to-listen>>. Acesso em: 18 jan. 2022.

MOTTA, L. M. V. de M. M.; ROMEU FILHO, P. (Org.). Audiodescrição: transformando imagens em palavras. São Paulo: Secretaria de Estado dos Direitos da Pessoa com Deficiência, 2010.

N. Kaufmann & W. Prager: Wege zu Kraft und Schönheit (1925). Disponível em: <<https://www.youtube.com/watch?v=RsvR90zBV-w>>. Acesso em: 12 jan. 2022.

NAVES, S. B. et al. Guia para produções audiovisuais acessíveis. Brasília: Secretaria do Audiovisual/ Ministério da Cultura, 2016.

NEILSON, K.; MANVELL, R.; WOLLENBERG, H. H. The Penguin film review. Harmondsworth: Penguin, 1948.

NIELSEN, M. A. Neural Networks and Deep Learning. Disponível em: <<http://neuralnetworksanddeeplearning.com/chap2.html>>. Acesso em: 16 jan. 2022.

NOLETO, C. O que é Sistema Operacional? O guia completo! Disponível em: <<https://blog.betrybe.com/tecnologia/sistema-operacional-tudo-sobre/>>. Acesso em: 15 jan. 2022.

NORNES, M. Cinema babel : translating global cinema. p 21. Minneapolis; London: University Of Minnesota Press, Cop, 2007.

NORNES, M. Cinema babel : translating global cinema. p 25. Minneapolis ; London: University Of Minnesota Press, Cop, 2007.

O que é Inteligência artificial? Como funciona, exemplos e aplicações. Disponível em: <<https://www.totvs.com/blog/inovacoes/o-que-e-inteligencia-artificial/>>.

O que são drivers? Entenda o que são e como eles funcionam nos PCs . Disponível em: <<https://canaltech.com.br/produtos/drivers-entenda-o-que-sao-e-como-eles-funcionam-195604/#:~:text=De%20forma%20simples%2C%20o%20driver>>. Acesso em: 18 jan. 2022.

Off-Screen Cinema Subtitling System. Disponível em: <<https://www.youtube.com/watch?v=b6uDWI4wVFg>>. Acesso em: 18 jan. 2022.

OPEN Source (Código Aberto): Como funciona e suas vantagens. Blog Totvs, 2021. Disponível em: <<https://www.totvs.com/blog/developers/como-funciona-um-software-open-source/>>. Acesso em: 04 de out. de 2021.

PÉREZ-GONZALEZ, Luis. Audiovisual Translation Theories, Methods And Issues. New York: Routledge, 2014.

RUGGLES, W. et al. The Plastic Age. Disponível em: <https://www.imdb.com/title/tt0016226/?ref_=nv_sr_1>. Acesso em: 13 jan. 2022.

São Paulo (Cidade). Secretaria Municipal da Pessoa com Deficiência - SMPED. Conhecer para Incluir a Pessoa com Deficiência. São Paulo: SMPED; Editora Mais Diferenças, 2020.

SAWAYA, Marcia Regina. Dicionário de informática e internet. São Paulo: Nobel, 1999.

SCHAPIRE, R. E. The strength of weak learnability. 30th Annual Symposium on Foundations of Computer Science, 1989.

Silent Era: Progressive Silent Film List. Disponível em: <<http://www.silentera.com/PSFL/data/C/ChamberMystery1920.html>>. Acesso em: 12 jan. 2022.

Sir Howard Stringer, Sony Chairman and CEO, Presents “The CES Seven,” Imperatives for Creating the Ultimate User Experience, at 2009 Consumer Electronics Show. Disponível em: <https://www.sony.com/content/sony/en/en_us/SCA/company-news/press-releases/sony-electronics/2009/sir-howard-stringer-sony-chairman-and-ceo-presents-the-ces-seven-imperatives-for-creating-the-ultimate-user-experience-at-2009-consumer-electronics-show.html>. Acesso em: 18 jan. 2022.

SLIDE, A. The new historical dictionary of the American film industry. Lanham, Md. London: Scarecrow, 2001.

Step by Step: Alpha Beta Pruning. Disponível em: <<https://www.youtube.com/watch?v=xBXHz4Gbdo>>. Acesso em: 15 jan. 2022.

STERNBERG, J. VON et al. Morocco. Disponível em: <<https://www.imdb.com/title/tt0021156/>>. Acesso em: 13 jan. 2022.

Subtitling for Cinema: A Brief History | The Artifice. Disponível em: <<https://the-artifice.com/subtitling-cinema-history/>>. Acesso em: 12 jan. 2022.

THE BENSHI TRADITION: CINEMA=PERFORMANCE. Disponível em: <<http://www.altx.com/interzones/kino2/benshi.html>>. Acesso em: 12 jan. 2022.

The British Film Catalogue. 1895-1970. Londres. Davi & Carlos. ISBN 0-715305572-4. Disponível em: <<http://www.britishpictures.com/books/catalogue.htm>> Acesso em: 5 jan. 2022.

The first movie with verbal texting - “College Chums” 1907 by Edwin S Porter Thomas Edison. Disponível em: <<https://www.youtube.com/watch?v=5JS4rLQtmA>>. Acesso em: 12 jan. 2022.

The Invisible Subtitler - A Documentary (SDH Subtitles included). Disponível em: <<https://www.youtube.com/watch?v=Pz75i6EsOto>>. Acesso em: 5 jan. 2022.

The Rank Organization. Disponível em: <https://en.wikipedia.org/wiki/Rank_Organisation>. Acesso em: 13 jan. 2022.

TIOBE Index | TIOBE - The Software Quality Company. Disponível em: <<https://www.tiobe.com/tiobe-index/>>. Acesso em: 18 jan. 2022.

Transcribe.Wreally - Digital Transcription Application Demonstration, Video 1. Disponível em: <<https://www.youtube.com/watch?v=TsUflTTOvu4>>. Acesso em: 18 jan. 2022.

TRANSLATOR. The New York Times, 8 nov. 1983.

TRINT. Realtime Transcription by TrintYouTube, 12 nov. 2019. Disponível em: <<https://www.youtube.com/watch?v=90iWFDxjF-c>>. Acesso em: 18 jan. 2022.

User's Guide. [s.l: s.n.]. Disponível em: <<https://files.support.epson.com/docid/cpd4/cpd40542.pdf>>. Acesso em: 18 jan. 2022.

Ways to Strength and Beauty. Disponível em: <https://en.wikipedia.org/wiki/Ways_to_Strength_and_Beauty>. Acesso em: 12 jan. 2022.

WEINBERG, H. G. A Manhattan odyssey : a memoir. New York: Anthology Film Archives, 1982. p108.

What is a Stenographer and How Does Stenography Work? | Naegeli. Disponível em: <<https://www.naegeliusa.com/blog/what-is-a-stenographer>>. Acesso em: 13 jan. 2022.