

A utilização de ferramentas de inteligência artificial no uso de traduções de vídeos para linguagens estrangeiras: uma abordagem Técnica e Ética

Yago Matheus de França Barata

Faculdade de Computação – Universidade Federal do Pará (UFPA)
Av. dos Universitários - Jaderlândia – 68746-360 – Castanhal – PA – Brasil

yagofranca998@gmail.com

***Abstract.** This paper examines the use of AI tools in dubbing and translating audiovisual content, considering technical and ethical aspects. With the growth of streaming and the globalization of productions, the demand for automated dubbing has increased. The tools Rask, Notta AI Showcase, and Vidnoz were analyzed, highlighting efficiency, cost reduction, and accessibility. However, ethical dilemmas arise, such as the devaluation of human labor and the protection of image rights. It is concluded that, despite the advances, it is essential to regulate AI usage to balance technological progress with human artistic value.*

***Resumo.** Este artigo examina o uso de ferramentas de IA na dublagem e tradução de conteúdo audiovisual, considerando aspectos técnicos e éticos. Com o crescimento do streaming e a globalização de produções, a demanda por dublagem automatizada aumentou. Foram analisadas as ferramentas Rask, Notta AI Showcase e Vidnoz, destacando eficiência, redução de custos e acessibilidade. Contudo, surgem dilemas éticos como a desvalorização do trabalho humano e a proteção dos direitos de imagem. Conclui-se que, apesar dos avanços, é essencial regulamentar o uso de IA para equilibrar progresso tecnológico e valor artístico humano.*

1. Introdução

Neste artigo, abordaremos sobre a utilização de Inteligência Artificial na dublagem e tradução e na apresentação de aplicações nesse âmbito, considerando seus prós e contras técnicos e também possível impactos éticos para a classe de trabalhadores da Dublagem.

Com o advento dos serviços de streamings e produções tanto no cenário cinematográfico quanto na produção de jogos e outras formas de entretenimento, a dublagem se tornou veementemente necessária já que grande parte destes não são produzidos no Brasil. É válido pontuar a importância cultural de tornar produções audiovisuais compreensíveis a sua população. Doria (2019) diz que analfabetos, deficientes visuais, crianças ainda não alfabetizadas, idosos e pessoas com demais dificuldades de leitura não têm condições de acessarem tais produtos por meio da legenda, como também, pouca parte da população possui níveis avançados de proficiência em línguas estrangeiras diversas para assistirem a filmes, séries, animações, etc. com o áudio em suas línguas originais. E com isso, podemos afirmar toda importância cultural referente a dublagem brasileira e sua aplicação na sociedade brasileira abrange quase que a totalidade da sociedade.

Outro ponto que será abordado neste artigo é com o avanço e popularização do uso da Inteligência Artificial e como este fenômeno vem incidindo na dublagem, quais os dilemas éticos e morais, e também no uso indevido de tal ferramenta no aspecto de produções audiovisuais. Hoje, é comum a presença de várias ferramentas gratuitas e de fácil acesso para o usuário comum na internet, dentre essas existem aplicações de IA específicas para manipulação e criação de Voz que podem ser usadas para criação de produções audiovisuais utilizando vozes de pessoas reais e de personagens fictícios para descrever uma cena ou ler certo texto. De acordo com uma pesquisa realizada pela Folha, 6 a cada 10 brasileiros preferem conteúdos dublados. Isso não está presente apenas no cinema, a Netflix divulgou recentemente cerca de 85% do seu público preferem programas e séries com dublagem em português.

Objetivo Geral

Apresentar ferramentas de IAs que trabalhem com dublagem de voz e conversão de áudio em línguas estrangeiras, destacando sua aplicabilidade e seus impactos técnicos e éticos.

Objetivos Específicos

Avaliar prós e contras nos âmbitos éticos e técnicos de determinadas aplicações;

Demonstrar cenários de aplicabilidade do tema do artigo;

Sugerir modelos híbridos para pesquisas futuras;

2. Referencial Teórico

A história da inteligência artificial (IA) e da dublagem oferece uma visão intrigante da convergência entre a arte da comunicação e a tecnologia. Como prática cultural, a dublagem é essencial para a criação da identidade cultural e a disseminação de narrativas internacionais, além de proporcionar acesso a obras audiovisuais em múltiplas línguas.

Por outro lado, a evolução da inteligência artificial, especialmente nas últimas décadas, tem transformado a forma como as vozes são sintetizadas e a maneira como as traduções são realizadas, levantando questões sobre a autenticidade, a qualidade e o impacto emocional das obras dubladas. Abordar essa história é essencial para compreendermos não apenas o desenvolvimento das técnicas de dublagem, mas também as implicações éticas e sociais da integração da IA nesse processo, permitindo uma reflexão crítica sobre o futuro da comunicação audiovisual e sua relação com a tecnologia.

2.1. A História da Dublagem e da Inteligência Artificial

A prática da dublagem teve início em 1927 com a chegada dos primeiros filmes sonoros, que substituíram os filmes mudos onde os diálogos, quando existiam, eram exibidos em quadros após a movimentação dos lábios dos personagens. Foi nesse período que surgiram as primeiras legendas, permitindo que pessoas de diferentes países compreendessem o enredo dos filmes. Da mesma forma, surgiram também os primeiros filmes dublados, com o objetivo de facilitar a compreensão universal das histórias cinematográficas.

A história da dublagem no Brasil inicia-se em 1930 com a chegada de animações infantis. Porém, o evento considerado o marco da dublagem no país ocorre em 1937, onde os cantores Carlos Galhardo e Dalva de Oliveira foram convidados para dar voz aos personagens do filme de animação da Disney “Branca de Neve e os Sete Anões”. Na época, todas as produções que eram produzidas no exterior apresentavam legendas.

Nos anos 60, com a popularização de televisores, um dos presidentes da época, Jânio Quadros determinou que filmes e outras produções audiovisuais estrangeiras fossem dubladas, motivado principalmente pela qualidade da imagem dos televisores da época não possuírem qualidade que tornasse legível o uso de legendas. Já no ano de 2013 ocorreu a estreia do Voice Arts Awards, conhecido como o principal prêmio da dublagem, inspirado no formato do Oscar. Este importante evento foi idealizado pela SOVAS (Sociedade de Artes e Ciências da Voz) com o propósito de celebrar globalmente os notáveis talentos envolvidos na arte da dublagem.

O marco zero da Inteligência Artificial se dá no ano de 1956, onde ocorreu a Conferência do Colégio de Dartmouth em New Hampshire (USA), onde o termo “Artificial Intelligence” foi apontado pela primeira vez, apresentando-se como um novo campo do conhecimento tecnológico (RUSSEL; NORVIG, 2009). Contudo, ideias que anteriores a esse evento já concatenavam com o conceito de inteligência artificial, sendo a maior delas o Teste de Turing, realizado por Alan Turing, matemático e cientista da computação britânico.

Em 1950, Turing publicou um artigo que pode ser dividido em duas partes, na primeira parte é apresentado o conceito do “Jogo da Imitação”. Nesse jogo, três participantes – um homem, uma mulher e um juiz (que pode ser homem ou mulher) – são colocados em salas isoladas e podem se comunicar exclusivamente por meio de mensagens digitadas. O objetivo do homem e da mulher é enganar o juiz, fazendo-se passar, respectivamente, por mulher e homem. Posteriormente, Turing substituiu um dos humanos por um computador, dando origem ao que ficou conhecido como Teste de Turing. Neste teste, um ser humano, um computador e um interrogador humano (juiz) são mantidos em salas distintas, podendo novamente se comunicar apenas por texto impresso. O ser humano e a máquina interagem entre si, enquanto o juiz deve analisar as respostas e tentar discernir qual é o computador e qual é o ser humano. A indagação que Turing formulou foi se a máquina poderia replicar o pensamento humano a ponto de confundir o juiz.

2.2. Inteligência Artificial e Aprendizado de Máquinas

Na tentativa de definir o que pode ser considerado conceitualmente como Inteligência Artificial notamos a dificuldade de definir em uma definição só, mas com o passar do tempo ela surgiu por quatro linhas de pensamentos, elas são:

I.Sistemas que pensam como seres humanos: “O novo e interessante esforço para fazer os computadores pensarem... máquinas com mentes, no sentido total e literal”. (HAUGELAND, 1985).

II.Sistemas que atuam como seres humanos: “A arte de criar máquinas que executam funções que exigem inteligência quando executadas por pessoas.” (KURZWEIL, 1990).

III.Sistemas que pensam racionalmente: “O estudo das faculdades mentais pelo seu uso de modelos computacionais.” (CHARNIAK; MCDERMOTT, 1985).

IV.Sistemas que atuam racionalmente: “A Inteligência Computacional é o estudo do projeto de agentes inteligentes.” (POOLE et al., 1998)

As linhas I e III referênciam a capacidade da IA processar pensamento e raciocínio, já II e IV relacionam-se com o comportamento. Adicionalmente, I e II denotam a fidelidade ao desempenho humano, enquanto III e IV mensuram o êxito do sistema se comparado a um conceito ideal de inteligência artificial. Pode-se dizer que um sistema é racional quando ele “faz o esperado”, com os dados que se tem. (RUSSELL; NORVIG, 2009).

Ou seja, podemos resumir como inteligência artificial, o campo da ciência da computação que se dedica à criação de sistemas capazes de realizar tarefas que, normalmente, exigiriam inteligência humana. Isso inclui a capacidade de aprender com experiências, raciocinar, resolver problemas, compreender a linguagem natural e reconhecer padrões. Pode ser dividida em duas categorias principais: a IA fraca, que é projetada para realizar tarefas específicas, e a IA forte, que busca replicar a inteligência humana de maneira geral.

Sobre o aprendizado de máquina, podemos defini-la como uma ramificação da inteligência artificial que permite que sistemas aprendam e se aprimorem automaticamente a partir de dados, sem serem explicitamente programados para isso. Os algoritmos são os motores que alimentam o machine learning. Em geral, dois tipos principais de algoritmos de machine learning são usados hoje: aprendizado supervisionado e aprendizado não supervisionado. A diferença entre eles é definida por como cada um aprende sobre dados para fazer previsões.

No ML Supervisionado, cientistas de dados atuam como guias e ensinam os algoritmos quais rumos devem ser tomados baseado nos dados. O algoritmo é treinado por um conjunto de dados que já está rotulado e tem uma saída predefinida. Já no ML não Supervisionado adota uma abordagem mais autônoma, na qual um sistema computacional é capaz de identificar processos e padrões complexos sem a necessidade de orientação contínua e próxima por parte de um ser humano. Esse tipo de aprendizado é caracterizado por um treinamento que se baseia em dados desprovidos de rótulos ou de uma saída específica previamente estabelecida.

Outra ramificação importante da inteligência artificial é o Processamento de Linguagem Natural (NLP), que permite que máquinas compreendam, interpretem e gerem linguagem humana. Ele envolve etapas como tokenização, análise sintática e semântica, e modelagem de linguagem. As aplicações incluem assistentes virtuais, tradução automática, chatbots, análise de sentimentos em redes sociais e resumo automático de textos. Essas ferramentas ajudam a facilitar a interação entre humanos e computadores de forma mais natural.

2.3. Tecnologias de IA com Áudio

Sobre as tecnologias atuais da inteligência artificial na dublagem, temos o TTS (Text-to-Speech) que é um método que converte texto escrito em fala audível. Inicialmente, temos como entrada, contendo símbolos, abreviações e sinais. O texto é posto em uma análise para composição fonética, geralmente separado em sílabas para melhor análise

da pronúncia. Com posse do texto fonético e tendo acesso as regras fonéticas, obtêm-se um conjunto de parâmetros de entrada e da variação deles no decorrer da sentença e através disto o sintetizador pode produzir uma fala contínua. (EGASHIRA, 1992).

O TTS é a principal forma que a inteligência artificial é capaz de sintetizar a voz humana, ou seja, a capacidade da IA produzir artificialmente a fala. Pode ser criada ao concatenar diferentes pedaços de fala gravada, quanto maior o pedaço da fala armazenado, maior é a qualidade e clareza do áudio obtido, como também maior o custo de armazenamento em um banco de dados fonético.

Esta síntese de fala pode ser, em resumo, concatenativa ou agrupamento de formantes. Na concatenativa, podemos dividi-la em três tipos básicos:

Por seleção de unidades onde utiliza-se uma base de dados com voz gravada, dividindo a frase recebida em elementos menores (como fonemas, sílabas, palavras, frases e orações) e com um apoio de um aparelho reconhecedor de voz modificado para compelir o alinhamento com o texto que será falado pela inteligência artificial. Tal sistema compara os elementos menores com a base de dados disponível analisando aspectos como forma de onda, espectrograma, frequência, duração do elemento e sua relação com elementos vizinhos. Geralmente, este método de síntese de fala utiliza uma árvore de decisão ponderada para criação da fala resultante. Como depende majoritariamente da robustez e fidelidade do banco de dados utilizado, pode obter resultados de falas que soem pouco naturais.

Por síntese de dífonos, que utiliza a base de dados mínima com todos os dífonos presentes em certa língua. Um dífono neste caso corresponde ao menor elemento capaz soar diferente dos demais, podendo ser uma consoante, vogal, fonema, entonação, dígrafos e outros. A partir deste banco de dados, reduz-se a frase de entrada no menor elemento possível e é alinhado ao banco. Tal modelo de síntese, em sua maioria, resulta em uma qualidade inferior de fidelidade por utilizar elementos mínimos.

Já pela síntese específica de domínio, concatena palavras e frases gravadas para criar saídas completas. Contudo, é limitada a um domínio específico onde a variedade de textos está limitada pela aplicação, por exemplo, sínteses de fala específicas de domínio são usadas em GPS e informativos de metrô. Apesar de sua simplicidade foi a pioneira comercial da popularização do uso da síntese de voz. A naturalidade da fala obtida pode ser muito fiel a realidade, já que a quantidade de variações de orações é limitada.

Outro método de síntese de voz, caracterizado por não utilizar a concatenação de elementos da fala, é a síntese de formantes. Este meio não depende de um banco de dados de amostras de falas humanas para produção de voz artificial, pois a fala resultante é obtida através de um modelo acústico artificial. Vários elementos da voz como frequência fundamental, níveis de ruído e pitches variam com o tempo para gerar uma forma de onda artificial. Neste modelo utiliza-se a análise morfológica da vibração de elementos da fala para síntese da voz. Muitos sistemas que utilizam este método apresentam resultados com fala robótica, porém compensa na inteligibilidade do resultado obtido e em sua alta velocidade de processamento.

Uma tecnologia atual que é válida salientar são os softwares de reconhecimento de fala que são sistemas ativados por voz, projetado para permitir que computadores e outros dispositivos interpretem a fala humana e depois a transcrevam em texto. Como

maior exemplo temos as plataformas de assistente virtual como Siri, Alexa e Google Assistente. Um programa ou aplicativo captura a fala do usuário e particiona o áudio em segmentos. Estes são analisados por meio de algoritmos de reconhecimento da fala humana para identificar que palavras melhor representam cada som. Então, o áudio é transcrito de forma inteligível ao sistema.

O processo que os softwares passam para assimilar a linguagem oral e transformá-la em informação digital é muito parecido com a aprendizagem humana. Geralmente captura-se ondas sonoras da fala humana e as convertendo em sinal digital. O sinal é particionado em pequenas partes de poucos milissegundos para facilitar a análise. Assim como um bebê capta primeiro sílabas e as replica de forma similar esses softwares aprendem como reconhecer a alocação. Cada parte deste sinal digital que anteriormente era uma onda sonora é comparado a um extenso banco de dados de sons conhecidos, fonemas, palavras e até pequenas frases. Modelos estatísticos como os de Markov, são usados para cálculo probabilístico de uma sequência específica de sons soar como uma palavra ou frase.

Recentemente, com o advento do uso do Deep Learning têm sido usadas para melhorar a precisão do reconhecimento de voz. Tais redes são capazes de distinguir e aprender padrões complexos como sotaques, ambientes ruidosos, entonação de voz que demonstre um sentimento com base num banco de dados potente. Também tem sido objeto de estudo reconhecimento de fala através de modelos com poucos dados é possível realizar o reconhecimento da fala humana através do Wav2vec 2.0, uma arquitetura robusta baseada em Deep Learning de um modelo pré-treinado de forma auto-supervisionada no domínio de tempo para conseguir representações discretas da fala. Um encoder convolucional trabalha em várias camadas do áudio bruto para representações da fala em tempos delimitados. Algumas dessas representações são mascaradas à rede, com o objetivo de fazer com que o modelo seja capaz de prever as partes mascaradas. (BAEVSKI, 2020)

3. Abordagem Ética

O uso da inteligência artificial na dublagem levanta questões éticas significativas, especialmente no que diz respeito à desvalorização do trabalho humano. Empresas, buscando reduzir custos operacionais, podem optar por soluções automatizadas mais baratas, o que pode diminuir a demanda por dubladores profissionais e afetar negativamente suas carreiras. No entanto, dublar vai além de reproduzir uma voz humana; envolve aspectos culturais, como sotaques, gírias, improvisações e interpretações emocionais, que são difíceis de replicar de forma satisfatória por sistemas de IA, por mais avançados que sejam.

Outro aspecto crucial envolve a propriedade intelectual e os direitos de imagem. A voz é um elemento fundamental na identidade de atores e cantores, e o uso inadequado da IA pode resultar em violações desses direitos, inclusive de artistas falecidos. Nesse contexto, torna-se essencial o estabelecimento de garantias legais, como contratos e acordos, que protejam o uso das vozes desses profissionais, garantindo que eles tenham controle sobre como suas vozes são utilizadas.

Além disso, a facilidade com que a IA pode criar vozes idênticas às dos artistas abre margem para manipulações e explorações indevidas. Essas vozes podem ser usadas em projetos que os artistas não aprovariam, ou até em contextos difamatórios, como a

disseminação de mensagens moralmente problemáticas e fake news. Isso reforça a necessidade de regulamentações para impedir o uso inadequado dessas tecnologias.

A diversidade e a representatividade também são preocupações pertinentes. A automação pode levar a uma padronização das vozes, especialmente quando há um número limitado de amostras sendo utilizadas, afetando negativamente a pluralidade linguística e cultural. A substituição de dubladores humanos pode reforçar estereótipos e minimizar a riqueza das expressões culturais. Além disso, futuros profissionais podem se sentir desestimulados diante da crescente preferência por soluções automatizadas, o que pode resultar em um mercado menos diversificado.

Por outro lado, a IA pode trazer benefícios importantes, como a democratização e ampliação da acessibilidade aos serviços de dublagem. Ao reduzir custos, a IA possibilita que produções com orçamentos menores tenham acesso a dublagem de alta qualidade, algo que antes seria inviável. Além disso, essas tecnologias podem oferecer soluções rápidas em situações de alta demanda ou com prazos apertados. No entanto, tais vantagens não devem ofuscar o valor artístico humano, que continua sendo insubstituível em termos de criatividade e expressão emocional.

Diante da complexidade do tema, é fundamental o desenvolvimento de regulamentações específicas para o uso da inteligência artificial na dublagem. Novos marcos legais devem assegurar a proteção dos direitos dos profissionais, garantir transparência no uso dessas tecnologias e informar adequadamente o público quando a dublagem for realizada por IA. Dessa forma, seria possível conciliar o avanço tecnológico com a preservação do trabalho artístico e da diversidade cultural.

4. Aplicabilidade

O emprego destas ferramentas de IA na dublagem estão revolucionando o processo de criação de vozes para personagens em filmes, séries e jogos. Utilizando técnicas avançadas de aprendizado de máquina, a IA pode analisar a voz original e replicar nuances emocionais, entonação e ritmo, permitindo a criação de dublagens realistas em várias línguas sem a necessidade de regravações extensas. Isso agiliza o processo e reduz custos, além de possibilitar a personalização de vozes para melhor representar diversidade e inclusão.

Imagine o seguinte cenário, o mais popular jogo de futebol da atualidade possui mais de 19000 futebolistas, segundo Leifert (2023) são necessárias mais de 80 mil frases para se realizar a narração de forma realista em um jogo de futebol. Para cada jogador que toca na bola há 3 entonações diferentes na voz dependendo de qual é a situação do jogo. Caso a bola esteja sendo tocada na defesa sem pressão do time adversário, o jogador que possui a bola terá seu nome falado pelo locutor de uma forma mais branda. Caso esteja criando uma jogada, seu nome será falado de uma forma mais enfática, e na possibilidade clara de gol, o locutor falará o nome do jogador de uma forma enérgica. Também há as falas “banter” que são falas aleatórias contando situações engraçadas do jogo ou do jogador para poder dar mais verossimilhança a narração do jogo. Temos também as locuções estatísticas de fatos que aconteceram no jogo e até em outras partidas, resultando em narrações mais realistas. Ou seja, um locutor teria que ter o trabalho de falar a mesma frase ou o mesmo nome várias vezes, e este retrabalho poderia ser auxiliado com o uso de inteligência artificial para facilitar a

qualidade do trabalho da narração do jogo, quanto a quantidade de falas do locutor em questão.

Outra conjuntura que pode ser aplicado o uso das aplicações IA de dublagem e tradução é a acessibilidade de conteúdos educacionais. Plataformas de e-learning e outras iniciativas de educação online não necessariamente estariam atreladas a linguagem que o tutor utilizou na produção do conteúdo. Ferramentas de IA podem capturar o áudio na língua nativa do tutor e transcrever em diversos idiomas, com a supervisão de profissionais da dublagem revisando e aprimorando o texto para melhor compreensão dos usuários de línguas estrangeiras. Ou seja, em vez de substituir o trabalho humano, a IA facilitaria a inclusão de mais idiomas tornando o processo de dublagem mais eficiente e também aumentando a abrangência de potenciais compradores do conteúdo de outro idioma.

Traduções e dublagens de projetos audiovisuais regionais e de baixo orçamento podem se beneficiar do uso destas aplicações, já que devido a possibilidade de tornar o conteúdo de fácil acesso para pessoas não nativas do idioma original pode ser interessante do ponto de vista comercial e financeiro. Por exemplo, filmes ou séries de sucesso em um país podem ser rapidamente adaptados para outros mercados com a ajuda da IA, enquanto os dubladores humanos revisam e ajustam as versões finais para garantir adequação cultural e expressiva. Sendo assim, a IA tem a capacidade de acelerar o processo de entrada dessas produções em mercados globais, tendo em vista que a supervisão de profissionais da dublagem na revisão final evita que traduções automáticas comprometam o significado cultural do conteúdo.

5. Análise de Aplicações

Neste capítulo, abordaremos diversas ferramentas de inteligência artificial com o enfoque em realizar a dublagem de voz em outras línguas diferentes do vídeo enviado para estudo da performance de cada plataforma, destacando suas principais características, prós e contras. A seguir, serão analisadas três ferramentas gratuitas de tratamento de dublagem e tradução audiovisual: Rask, Notta AI Showcase e Vidnoz, as quais oferecem soluções com diferentes benefícios e limitações. Para melhor mensuração dos nuances técnicos de tradução, em todas plataformas o vídeo base foi traduzido para inglês dos Estados Unidos.

O vídeo base é autoral, contém 15 segundos de duração e foi realizado com ruído de fundo para análise da capacidade do tratamento de áudio pela plataforma e verificação da dependência de qualidade. O conteúdo está disponível através do URL: https://drive.google.com/drive/folders/1dCZH4FvIPT-wxV1uVZ311Vg01BnDhsSH?usp=drive_link

5.1. Rask

Rask é uma ferramenta de IA voltada para a dublagem de vídeos em múltiplas línguas, projetada especificamente para a tradução de conteúdo audiovisual com vozes naturais. Ao acessar o site (<https://app.rask.ai>), é solicitado o upload de um vídeo onde informamos o nome do projeto, a língua original do vídeo e a língua para qual este será traduzido. Após o processamento do upload, o áudio é captado e segmentado em diferentes excertos que são traduzidos para o idioma solicitado conforme a imagem abaixo:

| Portuguese | English (US) | + New language |
|---|--|----------------|
| Olá, | Hello, | ... |
| tudo bem? | how are you? | ... |
| Meu nome é Iago | My name is Iago, | ... |
| e estou testando diferentes ferramentas | I am currently experimenting with various tools. | ... |
| que trabalham com dublagem | that work with dubbing | ... |
| com inteligência artificial. | using artificial intelligence. | ... |
| Muito obrigado pela atenção. | Thank you very much for your attention. | ... |

Figura 1. Subdivisão do áudio por tempo e tradução

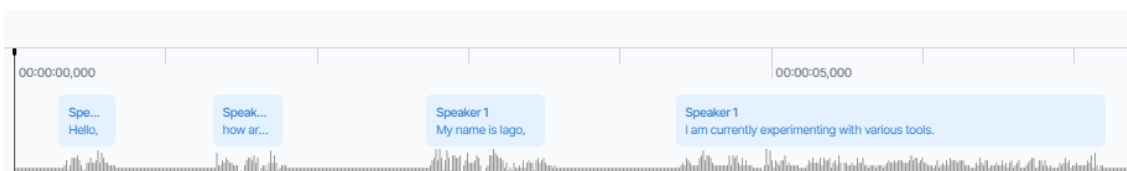


Figura 2. Posicionamento sincronizado da tradução do áudio base

A ferramenta oferece vozes sintetizadas com boa qualidade com base na qualidade do áudio enviado, que se aproximam da fala humana, podendo clonar a voz do vídeo base mesmo com poucos segundos de áudio. A interface do aplicativo é intuitiva permitindo que usuários com pouca experiência em dublagem e tradução possam utilizar a ferramenta sem problemas.

Também é capaz de analisar quando se fala e sincroniza com a voz gerada, e até permitindo alterações por segmento, porém não se preocupa com o tamanho da frase traduzida podendo destoar do vídeo ou gerar períodos de pausas perceptíveis ao usuário. Oferece uma dublagem automática excelente, mas carece de ajustes finos no controle de sotaque ou tom de voz, o que pode ser um empecilho para projetos que exigem uma entonação mais realista e gestão emocional.

5.2. Notta AI Showcase

Funciona de maneira semelhante ao aplicativo anterior, com os mesmos atributos solicitados, e após processamento do upload, o aplicativo já retorna com o vídeo traduzido em outro idioma. Acessível através do link: <https://www.notta.ai/en/showcase>

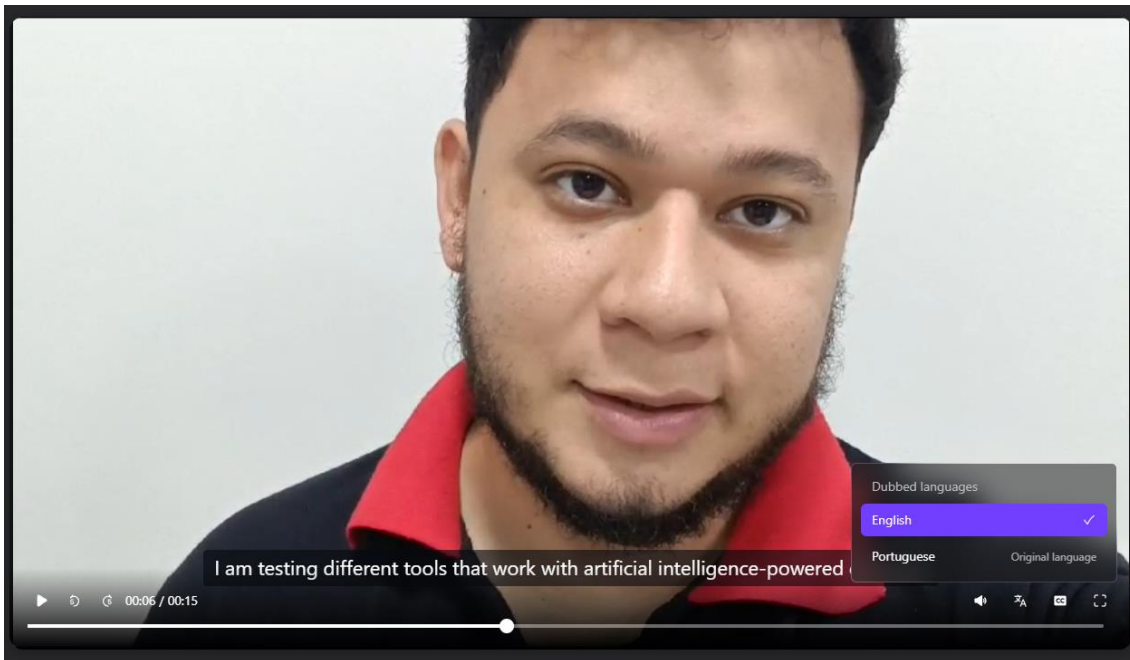


Figura 3. Tela obtida depois do upload do vídeo base

Os prós desta ferramenta se dão pela facilidade de uso com sua interface direta e pela integração com diversas plataformas podendo traduzir um vídeo com apenas o URL informado. Apesar de utilizar a própria voz do usuário, depende totalmente da qualidade do áudio enviado e de sua amostragem, por exemplo em casos de vídeos curtos, fica evidente a voz robotizada e a falta de sincronia com o movimento da boca do vídeo base. Não há possibilidade de alterar o tamanho do texto traduzido para poder acompanhar o vídeo, resultando em pausas perceptíveis.

5.3. Vidnoz

Vidnoz (<https://pt.vidnoz.com/>) é uma plataforma de inteligência artificial voltada para a geração automatizada de vídeos com dublagem multilíngue, focada em áreas como marketing e educação. Também é integrada com outras plataformas como Youtube, Drive e outros, permitindo o envio apenas do URL do vídeo fonte para traduções em diversas línguas. Neste aplicativo, temos uma tecnologia de sincronização labial que torna muito realista o vídeo traduzido e acaba sendo o diferencial desta plataforma.

Contudo, não permite ao usuário a modificação da tradução realizada pela própria ferramenta, e apesar de usar a voz do próprio usuário no vídeo base para síntese da voz usada no resultado, depende exclusivamente da qualidade do áudio e de sua amostragem resultando em voz genérica perceptível em vídeos curtos.

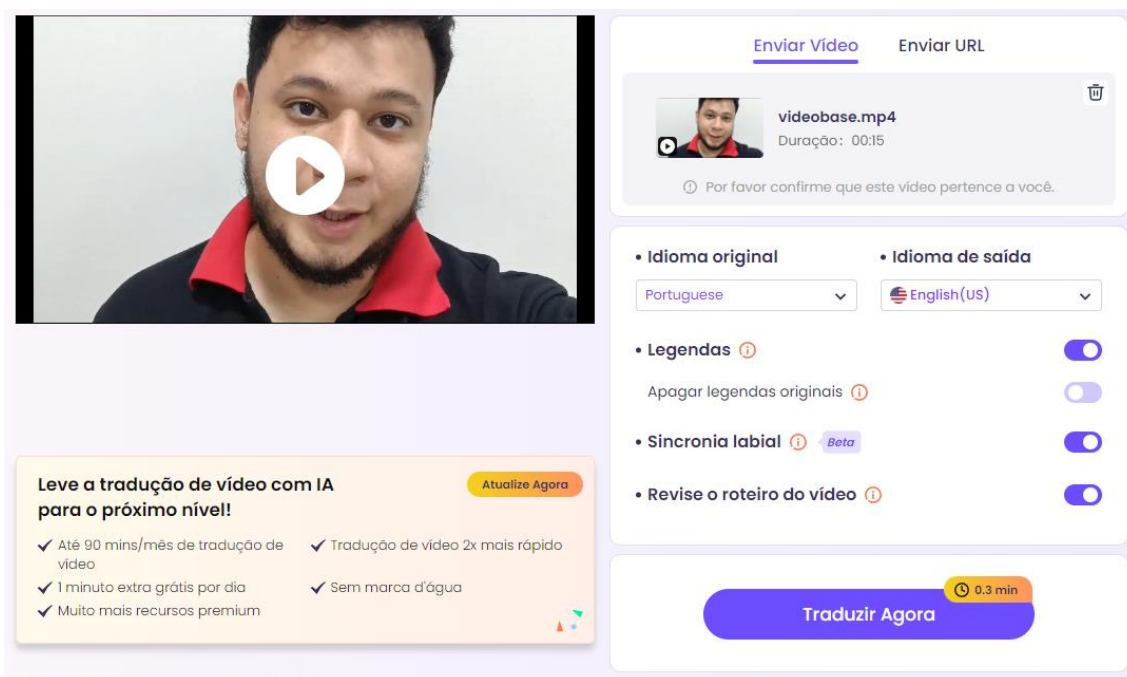


Figure 4. Interface simples e direta do Vidnoz

Vidnoz destaca-se pelo custo-benefício e pela agilidade em produções de pequeno e médio porte, porém, carece de controles avançados para dublagens mais complexas, tornando-a mais adequada para projetos com demandas técnicas menos rigorosas.

6. Conclusão

Este trabalho teve como objetivo central do trabalho discorrer sobre o uso de Inteligência Artificial na Dublagem, seu funcionamento e aplicabilidades em diversos setores do mercado atual, destacando seus impactos éticos e suas particularidades técnicas.

Ao longo da pesquisa realizada, analisando ferramentas de IA, como Rask, Notta AI Showcase e Vidnoz, demonstram diversas vantagens e também limitações na dublagem e tradução de conteúdos audiovisuais. Tendo isso em vista, podemos demonstrar como maior vantagem a possibilidade de usuários com pouca familiaridade com dublagem e tradução possam ter seus conteúdos de interesse traduzidos para diversos idiomas, com redução de custos e maior eficiência. Sob outra perspectiva, surgem dilemas éticos, principalmente relacionados à desvalorização do trabalho humano e à proteção dos direitos de imagem.

Tais achados apontam que, apesar que o uso dessas aplicações de IA possa trazer vantagens operacionais e econômicas para o setor de dublagem, faz-se necessário regulamentações do uso do serviço para que seja possível a preservação do valor artístico humano. Por exemplo a ferramenta Vidnoz, que é capaz de gerar dublagens e traduções automatizadas com rapidez, podem ser benéficas para projetos com prazo final curto e de baixo orçamento, contudo não substituem a autenticidade de dubladores profissionais e de sua expressividade.

É válido salientar que o estudo se conteve em abordar a análise teórica e a prática limitada das ferramentas em planos gratuitos, não tendo enfoque na recepção do público final em relação ao desempenho das ferramentas em gerar um resultado verossímil. Estudos futuros poderiam considerar investigar a percepção dos espectadores diante do resultado de vídeos processados de outras línguas para seu idioma nativo como também sua naturalidade a respeito das vozes sintetizadas.

Tendo em vista o iminente avanço de novas tecnologias de inteligência artificial, pesquisas futuras poderiam analisar a viabilidade de modelos híbridos tendo a supervisão criativa de dubladores profissionais sobre a automação de ferramentas de IA visando maximizar a qualidade das dublagens e mantendo sua originalidade artística.

Desta forma, conclui-se que a utilização de aplicações de dublagem que utilizam inteligência artificial concebem significativo avanço na democratização do acesso a conteúdos audiovisuais. Todavia, para que esta evolução ocorra de forma ética e benéfica, é fundamental a criação de regulamentações que protejam o valor do profissional de dublagem, considerando o caráter insubstituível de sua criatividade e sensibilidade.

References

- FOLHA DE S.PAULO. Pesquisa revela que 6 em 10 brasileiros preferem filmes dublados. Folha de S.Paulo, São Paulo, 03 ago. 2015. Disponível em: <https://m.folha.uol.com.br/ilustrada/2015/08/1666126-pesquisa-revela-que-6-em-10-brasileiros-preferem-filmes-dublados.shtml>. Acesso em: 01 ago. 2024.
- DORIA, Nívea Guimarães. Revisão Bibliográfica Acadêmica Sistemática sobre Dublagem Brasileira em uma Pesquisa de Doutorado. Palimpsesto-Revista do Programa de Pós-Graduação em Letras da UERJ, v. 18, n. 29, p. 362-378, 2019.
- SUPER INTERESSANTE. Quando surgiu a dublagem no Brasil e no mundo? Super Interessante, São Paulo, 22 out. 2024. Disponível em: <https://super.abril.com.br/mundo-estranho/quando-surgiu-a-dublagem-no-brasil-e-no-mundo>. Acesso em: 02 ago. 2024.
- DE CASTRO BARBOSA, Xênia. Breve introdução à história da Inteligência Artificial. Jamaxi, v. 4, n. 1, 2020.
- RUSSELL, Stuart J.; NORVIG, Peter. Artificial intelligence: a modern approach. Pearson, 2016.
- ONODY, Roberto N. Teste de Turing e Inteligência Artificial. IFSC (Instituto de Física de São Carlos) – USP, 28 set. 2021. Disponível em: <https://www2.ifsc.usp.br/portal-ifsc/teste-de-turing-e-inteligencia-artificial/>. Acesso em: 01 ago. 2024.
- HAUGELAND, John. Artificial Intelligence: The Very Idea. Massachusetts: The MIT Press, 1985.
- KURZWEIL, Ray. The Age of Spiritual Machines. Massachusetts: The MIT Press, 1999.
- CHARNIAK, Eugene; MCDERMOTT, Drew. A Bayesian Model of Plan Recognition. Massachusetts: Addison-Wesley, 1985.

- POOLE, D.; MACKWORTH, A. K.; GOEBEL, R. Computational Intelligence: A Logical Approach. Oxford: Oxford University Press, 1998.
- EGASHIRA, Francisco. Síntese de voz a partir de texto para a Língua Portuguesa. 1992. Dissertação (Mestrado) – Universidade Estadual de Campinas, Campinas, 1992.
- RIBEIRO, Jussara et al. Síntese computacional de voz. 2010. Tese (Doutorado) – Universidade de São Paulo, São Paulo, 2010.
- BAEVSKI, Alexei et al. wav2vec 2.0: A framework for self-supervised learning of speech representations. Advances in neural information processing systems, v. 33, p. 12449-12460, 2020.
- LEIFERT, Tiago. TIAGO LEIFERT - Flow #268. Apresentador Igor Rodrigues Coelho. Podcast: Flow Podcast. São Paulo: Estudos Flow, 2023. Disponível em: <https://www.youtube.com/watch?v=-roK1JndcBY>. Acesso em: 13 ago. 2024.
- RASK. Rask: Ferramenta de Tradução e Dublagem de Vídeos. Disponível em: <https://app.rask.ai>. Acesso em: 10 out. 2024.
- NOTTA AI. Notta AI Showcase: Ferramenta de Dublagem Multilíngue. Disponível em: <https://www.notta.ai/en/showcase>. Acesso em: 10 out. 2024.
- VIDNOZ. Vidnoz: Plataforma de Dublagem Automatizada. Disponível em: <https://pt.vidnoz.com>. Acesso em: 10 out. 2024.