



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

**RONNYLDO RODRIGUES ELOY DA SILVA**

**INTERNACIONALIZAÇÃO DE WEBSITES**

**CAMPINA GRANDE - PB**

**2023**

**RONNYLDO RODRIGUES ELOY DA SILVA**

**INTERNACIONALIZAÇÃO DE WEBSITES**

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

**Orientador : Eanes Torres Pereira**

**CAMPINA GRANDE - PB**

**2023**

**RONNYLDO RODRIGUES ELOY DA SILVA**

**INTERNACIONALIZAÇÃO DE WEBSITES**

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

**BANCA EXAMINADORA:**

**Eanes Torres Pereira**

**Orientador – UASC/CEEI/UFCG**

**Franklin de Souza Ramalho**

**Examinador – UASC/CEEI/UFCG**

**Francisco Vilar Brasileiro**

**Professor da Disciplina TCC – UASC/CEEI/UFCG**

**Trabalho aprovado em: 28 de Junho de 2023.**

**CAMPINA GRANDE - PB**

## RESUMO

A localização e internacionalização de websites são práticas que visam, entre outros objetivos, conquistar ou atrair clientes em novos mercados; e/ou conferir maior credibilidade e visibilidade à instituição, empresa ou produto a que se destina o seu conteúdo. Para atingir esse resultado, é necessária uma estratégia para determinar como o website será feito, pensando em cada público-alvo e região. É necessário ser capaz de suportar não apenas idiomas diferentes, mas regiões diferentes com o mesmo idioma. A interface do usuário precisa responder não apenas ao tamanho da tela, mas também a diferentes idiomas e modos de escrita. O conteúdo precisa ser estruturado, desde sua interface do usuário e o formato de suas datas, para ser adaptável a qualquer idioma que você lançar nele. Neste trabalho será apresentada a importância da internacionalização/localização de websites, explicando as principais etapas para a localização de um website, desde sua estrutura e design. Assim como funcionam os principais serviços de hospedagem de websites e seus serviços de tradução para solucionar a demanda de internacionalização. Por fim, foi desenvolvida uma aplicação web em Django e React, capaz de traduzir texto a partir de HTML, além de legendar e traduzir vídeos e áudios. A aplicação utiliza APIs da Google para o processo de tradução e legenda.

# **WEBSITE INTERNATIONALIZATION**

## **ABSTRACT**

The localization and internationalization of websites are practices that aim, among other objectives, to conquer or attract customers in new markets; and/or give greater credibility and visibility to the institution, company or product for which its content is intended. To achieve this result, a strategy is needed to determine how the website will be made, thinking about each target audience and region. You need to be able to support not just different languages, but different regions with the same language. The UI needs to respond not only to screen size, but also to different languages and writing modes. The content needs to be structured, from its UI and the format of its dates, to be adaptable to whatever language you throw at it. In this work, the importance of internationalization/localization of websites will be presented, explaining the main steps for the localization of a website, from its structure and design. This is how the main website hosting services and their translation services work to solve the demand for internationalization. Finally, a web application was developed in Django and React, capable of translating text from HTML, in addition to subtitling and translating videos and audios. The application uses Google APIs for the translation and subtitling process.

# Internacionalização de Websites

Aluno: Ronnyldo Rodrigues Eloy da Silva  
Unidade Acadêmica de Sistemas e Computação  
Universidade Federal de Campina Grande.  
Campina Grande, Paraíba - Brasil.  
ronnyldo.silva@ccc.ufcg.edu.br

Orientador: Eanes Torres Pereira  
Unidade Acadêmica de Sistemas e Computação  
Universidade Federal de Campina Grande.  
Campina Grande, Paraíba - Brasil.  
eanes@computacao.ufcg.edu.br

## RESUMO

A localização e internacionalização de websites são práticas que visam, entre outros objetivos, conquistar ou atrair clientes em novos mercados; e/ou conferir maior credibilidade e visibilidade à instituição, empresa ou produto a que se destina o seu conteúdo. Para atingir esse resultado, é necessária uma estratégia para determinar como o website será feito, pensando em cada público-alvo e região. É necessário ser capaz de suportar não apenas idiomas diferentes, mas regiões diferentes com o mesmo idioma. A interface do usuário precisa responder não apenas ao tamanho da tela, mas também a diferentes idiomas e modos de escrita. O conteúdo precisa ser estruturado, desde sua interface do usuário e o formato de suas datas, para ser adaptável a qualquer idioma que você lançar nele. Neste trabalho será apresentada a importância da internacionalização/localização de websites, explicando as principais etapas para a localização de um website, desde sua estrutura e design. Assim como funcionam os principais serviços de hospedagem de websites e seus serviços de tradução para solucionar a demanda de internacionalização. Por fim, foi desenvolvida uma aplicação web em Django e React, capaz de traduzir texto a partir de HTML, além de legendar e traduzir vídeos e áudios. A aplicação utiliza APIs da Google para o processo de tradução e legenda.

## Palavras-chave

Internacionalização de websites, localização, legendas de vídeo, legendas de áudio, aplicação web python.

## Repositório

<https://github.com/RonnyldoSilva/TCC>

## 1. INTRODUÇÃO

Este Trabalho de Conclusão de Curso trata da relação entre localização e tradução na web, sendo a primeira uma forma relativamente nova de serviço multilíngue. Tenta-se mostrar que alguns aspectos dos estudos de tradução são relevantes para a localização de websites e, vice-versa, a localização de websites pode fornecer alguns novos insights para a tradução.

Inicialmente, veremos alguns conceitos centrais da localização de websites para evitar mal-entendidos. A seguir, definiremos e descreveremos o processo de localização de websites e sua finalidade, concentrando-nos no objetivo geral de traduzir ou localizar websites. Este esclarecimento é seguido por uma breve descrição dos métodos e estratégias utilizadas no processo de localização de websites, incluindo o papel e a importância dos aspectos de localização relacionados aos negócios.

Um website [1] abrange todas as páginas da web acessíveis em um endereço comum da web (nome de domínio), como [www.google.com.br](http://www.google.com.br). Um website consiste em vários documentos, gráficos, programas e assim por diante, cada um dos quais é identificado por um identificador de recurso uniforme (URL). Em alemão, no entanto, há um problema de terminologia, pois não existe tal termo como website. Quando falamos de *Webseiten* (páginas da web) ou Homepage em alemão, queremos dizer todas as páginas da World Wide Web para uma determinada instituição ou empresa. Atualmente, o novo termo *Webauftritt* foi criado para o website em inglês. É importante observar que o objeto da localização é um website inteiro, e não apenas uma única página web.

O segundo termo usado na localização do título é um termo definido como a adaptação de um produto a um determinado local. Uma localidade (LISA 2003, Esselink 2001, Yanker 2003) [2][3][4] refere-se a um grupo de pessoas que compartilham um idioma, um sistema de escrita e outras propriedades que podem exigir uma versão separada de um produto. Pode ser uma região, um país ou apenas uma comunidade linguística.

A localização hoje é usada em conjunto com o termo internacionalização, que significa a preparação de um produto para torná-lo adequado para localização eficiente, e globalização significa o design global de um produto. A localização de websites é assim definida como o "processo de modificação de um website para uma localidade específica" (Yunker 2002) [4].

Como tradutores podemos nos perguntar se a tradução faz parte da localização, como apontado em Esselink (2000) onde o autor define a localização como a tarefa global da qual a tradução faz parte, assim como qualquer outra parte do processo como gerenciamento de projetos, adaptação de imagem ou configuração de um gateway de idioma. Isso implicaria que o profissional de localização está no comando e o tradutor atua como colaborador de textos em língua estrangeira. Visto de uma perspectiva de estudos de tradução, no entanto, o processo de tradução envolve linguística, bem como uma transferência cultural e a intenção comunicativa ou função do texto de chegada é de importância geral. A tradução, portanto, sempre envolve alguma forma de adaptação em relação ao próprio texto ou outros itens relevantes para o documento, como gráficos etc. A localização de websites coloca o texto em segundo plano e foca em aspectos multimídia. No entanto, o texto continua a ser uma informação chave dentro de uma página da web. A tradução, como tarefa, tem uma história centenária, enquanto a localização é um fenômeno dos últimos 20 anos, talvez apenas um novo nome, um tipo específico de tradução. Nesse caso, a tradução seria então o conceito mais amplo.

Na literatura (Resende 2009, Alves 2005, Darin, 2005) [14][15][16] há indicações para ambas as hipóteses: por um lado,

existem cursos de formação específicos para profissionais de localização, associações de localizadores profissionais para apoiar a noção de uma nova profissão de localização forte e, por outro lado, institutos de formação em tradução que oferecem cursos sobre localização e tradutores que trabalham na indústria de localização.

A seguir, serão descritas todas as etapas e desafios no desenvolvimento de websites, considerando os princípios da internacionalização e localização de websites. Assim como, o desenvolvimento de uma aplicação web capaz de auxiliar na tradução e legendamento do conteúdo de um website, focando no HTML, vídeo e áudio.

## 2. LOCALIZAÇÃO DO WEBSITE

Não existe uma experiência única de localização de websites. Cada projeto é tão diferente quanto seus idiomas e localidades, público-alvo, tipos de tradução e muito mais. Mas existem processos, tecnologias e práticas recomendadas comuns.

Neste trabalho, demonstraremos os estágios de um projeto típico de localização da Web, desde o planejamento estratégico até a entrada em operação de vários websites e além. Haverá perguntas a serem respondidas, escolhas a serem feitas e muito a ser aprendido ao longo deste documento.

### 2.1 PLANEJAMENTO E PREPARAÇÃO

Por meio de planejamento, preparação e estratégia de localização focada, você pode superar a complexidade para acelerar o tempo de colocação no mercado, conter custos e desfrutar de um investimento altamente lucrativo no futuro da sua empresa, instituição ou escola.

#### 2.1.1 O que é uma estratégia de localização de websites?

Uma estratégia de localização web deve ser vista como uma extensão do seu plano de negócios. Se você tem um negócio global ou está se tornando global, você precisa de websites e conteúdo consumíveis localmente. Sua estratégia de localização é seu plano para executar esse requisito de negócios.

Ao planejar seu conteúdo no idioma de origem, seja inglês ou outro, adote a mentalidade de que você também está planejando para suas necessidades de conteúdo global ou no idioma de destino. De fato, para evitar tratar a localização como uma reflexão tardia, uma estratégia de localização web eficiente não faz distinção entre os idiomas de origem e de destino.

Portanto, certifique-se de criar seu website com a globalização em mente. Porque quando você cria estratégias, planeja e usa os fluxos de trabalho e processos corretos, a localização pode ser relativamente perfeita. Mas, sem um planejamento adequado, pode ser um empecilho demorado e caro para o seu negócio.

#### 2.1.2 Algumas questões a considerar ao criar sua estratégia de localização.

O planejamento e a preparação são importantes para o seu projeto de localização. Só porque há um orçamento e um prazo não

significa que você está pronto. É melhor não adicionar complexidade tentando localizar, sem antes possuir uma estratégia claramente definida.

É importante fazer as seguintes perguntas antes de desenvolver seu website [3]:

- Quais são seus objetivos de negócios globais?
- Você definiu e analisou mercados-alvo, idiomas e localidades?
- Sua organização está pronta para se expandir internacionalmente?
- Quais são os objetivos comerciais de seu website corporativo em comparação com as versões globais?
- As principais partes interessadas concordam com os objetivos e o escopo?
- Você tem as tecnologias certas para oferecer suporte a uma operação multilingue na Web (especialmente um sistema de gerenciamento de conteúdo pronto para uso global)?
- Você construiu seu website com a globalização em mente?
- Você tem uma estratégia de otimização de mecanismo de busca multilingue e os recursos para executá-la?
- Você está preparado do ponto de vista operacional e de processos de negócios?
- O seu departamento de gestão ou financeiro está preparado para alocar um orçamento?
- Como seu projeto de localização afetará outros sistemas comerciais e de TI?
- Você registrou nomes de domínio com código de país?
- Sua empresa e marca trabalham em seus mercados-alvo?

#### 2.1.3 Por que devo localizar quando é mais fácil padronizar?

Algumas empresas escolhem uma estratégia de padronização para seus websites globais. Como o nome indica, isso envolve traduções simples do website principal da organização em vários locais. Os websites são espelhados, na verdade, com apenas um pequeno subconjunto de informações localizadas, talvez incluindo informações de contato, biografias executivas e comunicados à imprensa.

Uma estratégia de localização ou adaptação é exatamente o oposto da padronização. Oferece um mix de produtos e serviços adaptados a localidades específicas. Ele atende às necessidades locais e preferências culturais, ao mesmo tempo em que se conecta em um nível emocional para estimular a ação. E mostra respeito pelos consumidores do mercado-alvo, o que gera confiança e fortalece as marcas.

#### 2.1.4 Benefícios para localização

Com menos conteúdo para traduzir, os processos tradicionais de localização são simplificados, reduzindo o tempo de colocação no mercado, melhorando a qualidade do conteúdo e reduzindo os custos de localização. Os benefícios de uma abordagem baseada [3] em componentes para conteúdo global incluem:

- Traduzir o conteúdo uma vez: a localização inicial será a única vez em que um documento inteiro será traduzido. Ao usar componentes em nossa estratégia de

conteúdo, apenas o conteúdo atualizado ou modificado precisará ser retraduzido, não o documento inteiro. Uma vez atualizado o conteúdo, ele pode ser reaproveitado em toda a sua empresa, melhorando a qualidade e a consistência do seu conteúdo.

- Economize dinheiro na tradução de conteúdo: ao traduzir componentes em vez de documentos inteiros, você evita custos desnecessários de tradução e editoração eletrônica.
- Reduza o tempo de lançamento: traduzir menos palavras não apenas economiza seu dinheiro, mas também economiza tempo. Os tradutores não precisam perder tempo verificando o conteúdo já localizado, atrasando o tempo de lançamento no mercado. Além disso, ao criar componentes, você pode executar processos paralelos altamente eficientes nos quais localizam alguns componentes enquanto desenvolvem outros.

Então, uma vez que o conteúdo é componentizado e localizado, e você precisa atualizar o conteúdo, você apenas traduz novamente o conteúdo que foi modificado.

### 2.1.5 *Quais idiomas devem ser incluídos em sua estratégia?*

Os idiomas que você escolhe para seu projeto de localização dependem de suas metas de negócios globais, que são informadas por quais mercados, países ou localidades oferecem a maior oportunidade e o retorno do investimento mais promissor.

#### 2.1.5.1 *Múltiplos mercados domésticos?*

Lembre-se de que, em muitos países, os mercados domésticos têm populações grandes e crescentes de falantes não nativos. Nos Estados Unidos, por exemplo, é impossível ignorar a importância e o potencial dos mercados de língua espanhola.

#### 2.1.5.2 *O que os concorrentes estão fazendo?*

Além disso, preste muita atenção em quais mercados seus concorrentes estão buscando e em quais idiomas eles investiram. E considere a quantidade de tráfego do website que você já está recebendo de determinados países ou localidades.

#### 2.1.5.3 *Países multilíngues significam vários locais.*

Quantos idiomas são falados em seus países de mercado-alvo? Por exemplo, na Suíça, dependendo da localidade, é provável que você ouça alemão, francês, italiano ou romanche. Na Bélgica, é holandês (ou flamengo, sendo o holandês a língua raiz do dialeto flamengo), francês e alemão. E no Canadá, geralmente é necessário localizar para o francês e o inglês canadenses.

### 2.1.6 *Escrevendo para audiências globais*

Ao escrever para localização, mantenha suas frases simples para aumentar a compreensão do conteúdo de origem. Use um guia de estilo para consistência e escreva como se estivesse falando com alguém, sem jargões, gírias, clichês, coloquialismos ou termos culturais específicos. E lembre-se de que frases claras, concisas e bem construídas melhoram a qualidade das localizações no idioma de destino, reduzem o tempo de resposta e reduzem os custos. A

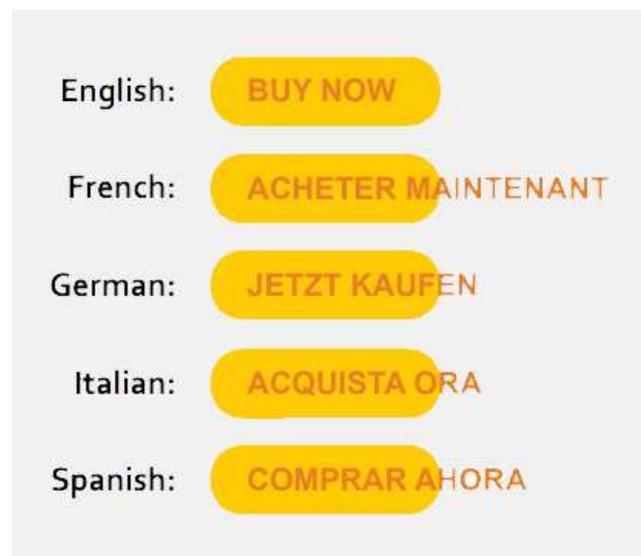
seguir, algumas dicas [4] para produzir textos mais facilmente traduzíveis:

1. Mantenha as frases breves. Para maior compreensão e traduções mais simples, devem conter no máximo 15 a 20 palavras.
2. Use a ordem de palavras do inglês padrão sempre que possível, o que geralmente significa um sujeito, verbo e objeto com modificadores associados.
3. Use apenas um termo para identificar um único conceito. Os sinônimos atrapalham a clareza.
4. Evite humor. Raramente traduz com equivalência.
5. Seja claro com as datas internacionais. Dependendo da localidade, em vez de usar “5/28/2023”, escreva o mês (28 de Maio, 2023). Use numerais para números grandes (milhões e bilhões) e converta e liste medições de peso, altura, largura, temperatura e valores monetários, etc. para cada par de idiomas. Preste atenção também à conversão local de hora, endereços, números de telefone, fins de semana, feriados e formatos numéricos.
6. Use a voz ativa em vez da passiva. É mais direto e fácil de entender e traduzir.
7. Deixe espaço suficiente para expansão de texto, principalmente em interfaces de software e gráficos. Prepare-se para um mínimo de 15%, com até 35% de expansão do texto original em inglês para alguns idiomas (alemão, por exemplo).

Embora os caracteres do Leste Asiático possam diminuir em 10% ou mais, a nuance e a complexidade dos traços em glifos individuais podem exigir um pequeno aumento no tamanho da fonte.

Quando traduzidos do inglês, alguns outros idiomas ocidentais podem expandir em até 35% ou mais. Portanto, o planejamento é fundamental, especialmente ao localizar interfaces de aplicativos da web, menus de navegação do website e outros elementos, como botões de chamada para ação, etc. A Figura 1 mostra um exemplo:

Figura 1: exemplo de planejamento para expansão de texto.



### 2.1.7 Criação de websites multilíngues

Há vários elementos importantes a serem considerados ao projetar websites para audiências internacionais. Aqui estão alguns dos mais importantes, juntamente com algumas práticas recomendadas [4].

#### 2.1.7.1 Imagens

A arte, por sua própria natureza, é subjetiva. Assim como gráficos comerciais, fotografias, ícones e símbolos, especialmente em websites globais. Designs que parecem inocentes ou elegantes para os ocidentais podem ser ofensivos em outras culturas. Evite imagens de mãos, animais ou símbolos religiosos. Use pessoas culturalmente apropriadas nas fotos. E jogue pelo seguro. Faça sua pesquisa e verifique com a equipe do país se você tem alguma reserva.

#### 2.1.7.2 Cor

O simbolismo interpretado das cores varia muito de cultura para cultura. Alguns especialistas afirmam que a cor azul é mais aceita globalmente. Talvez por causa de sua conexão com a natureza, tende a evocar sentimentos positivos e uma sensação de calma. No entanto, no Irã, esta cor está associada ao luto. Como as cores são tão subjetivas, é recomendável usá-las funcionalmente em seu website e pesquisar conflitos se você não tiver certeza das implicações culturais.

#### 2.1.7.3 Gráficos

Evite incorporar texto em gráficos. Isso complica o processo de tradução, ao impedir o acesso fácil ao texto e pode implicar na tarefa (evitável) de reconstruir arquivos com gráficos. Em vez disso, é sugerido que seus desenvolvedores sobreponham texto em gráficos dentro do HTML.

#### 2.1.7.4 PDFs e arquivos de origem

Os arquivos Adobe PDF [17] podem dificultar o acesso ao conteúdo para tradução. Os PDFs não podem ser editados em alguns programas e podem perder a formatação ao serem convertidos, gerando alterações no conteúdo, necessitando de edições, aumentando os custos e atrasos..

### 2.1.8 Escolhendo sua estrutura de URL/domínio multilíngue

Os visitantes do website e os mecanismos de pesquisa usam a estrutura de URL para tomar decisões sobre a relevância e autoridade dos websites para eles. De fato, as URLs fornecem contexto instantâneo para os visitantes, oferecendo informações sobre o conteúdo da página, e onde a página reside na hierarquia do website.

Ao tomar decisões sobre sua estrutura da URL, lembre-se de que (na medida do possível) as URLs devem identificar onde o conteúdo reside, qual o idioma do website, qual a finalidade, qual a localização e o público-alvo do conteúdo. A Figura 5 mostra um exemplo de estrutura da URL:

Figura 4: Estrutura da URL.



Como você estrutura suas URLs afeta diretamente o sucesso dos mecanismos de pesquisa que os usam para catalogar seus índices. Portanto, reservar um tempo para tornar suas URLs amigáveis aos mecanismos de pesquisa é um investimento importante, mas relativamente pequeno, no desempenho futuro de seus websites globais.

Existem três opções normalmente usadas para estruturar e otimizar URLs de websites multilíngues. Mas isso dependerá dos objetivos e recursos exclusivos da sua organização, empresa ou instituição, incluindo se está segmentando o público-alvo por país, idioma ou ambos. Por exemplo, essas abordagens podem direcionar o tráfego para o seu website alemão através do código do país, "de".

- Domínios de primeiro nível com código de país:
  - <https://www.mysite.de>
- Subdiretórios:
  - <https://www.mysite.com/de>
- Subdomínios:
  - <https://www.de.mysite.com>

Neste exemplo, para um website francês, basta substituir o código do país "fr".

## 2.2 HOSPEDAGEM DE WEBSITES

Um provedor de hospedagem de websites é um serviço que possibilita a pessoas ou empresas com sistemas online a guardar páginas, arquivos, informações, imagens, vídeos, ou qualquer conteúdo acessível pela Web. Provedores de hospedagem de websites tipicamente são empresas que fornecem um espaço em seus servidores e conexão à internet a estes dados aos seus clientes. Tipicamente, a hospedagem de websites hospeda arquivos 24 horas por dia, em um endereço IP estático, por sua vez, utilizamos um domínio para encontrar com facilidade este website, é muito comum confundir hospedagem de websites com registro de domínio, entretanto, apesar de ambos trabalharem em conjunto, são serviços diferentes (Wikipedia) [5].

Muitas empresas, instituições ou escolas utilizam de provedores de hospedagem para construir seus websites, no intuito de vender ou divulgar seus conteúdos. Diante disto, foram analisados os principais provedores de hospedagem, para verificar se os mesmos possuem alguma ferramenta para auxiliar na localização ou tradução de websites, assim, facilitando o processo de internacionalização de websites.

Na Tabela 1, estão listados os principais provedores de hospedagem internacionais. Na coluna “Provedor”, está o nome do provedor de hospedagem, assim o identificando. Na coluna “Tradução/Localização”, está a informação se o provedor possui algum meio de tradução ou localização. Por fim, na coluna “WordPress”, indica se o provedor suporta WordPress. Há vários plugins de tradução para o WordPress, entre os principais estão Weglot Translate, GTranslate, Polylang e TranslatePress.

Tabela 1:

Provedor	Tradução/Localização	WordPress
Hostinger	Não	Sim
GoDaddy	Não	Sim
Wix	Sim	Não
HostGator	Não	Sim
Weblink	Não	Sim
SiteGround	Não	Sim
1&1 IONOS	Não	Sim
King Host	Sim	Sim
UOL Host	Não	Sim
HostMídia	Não	Sim
LocaWeb	Sim	Sim
Metaweb	Não	Não
HostNet	Não	Sim

Apenas o Wix e King Host possuem ferramentas nativas para a tradução de websites. Com exceção da Wix e Metaweb, todos os principais provedores permitem a utilização de WordPress, assim permitindo de plugins de terceiros para a tradução do website.

### 3. APLICAÇÃO WEB

A aplicação Website Internationalization oferece uma solução para as necessidades de tradução de conteúdo de um website. A ideia principal do trabalho consiste no desenvolvimento de uma aplicação web, de uso simplificado e intuitivo, que permite que os usuários traduzam arquivos html, como também gerar legendas a partir de vídeos e áudios. Também é possível traduzir as legendas.

O desenvolvimento do sistema prezou pela responsividade, permitindo uma experiência agradável ao usuário independentemente do dispositivo usado para o acesso. O sistema deve permitir o carregamento de arquivos que serão traduzidos ou legendados. Retornando ao usuário um arquivo HTML traduzido,

ou um arquivo TXT contendo as legendas (traduzidas ou não) do vídeo ou áudio, anteriormente selecionado.

A aplicação possui um único tipo de usuário. O usuário é capaz de executar 3 ações: traduzir HTML, legendar vídeo e legendar áudio. Tais funcionalidades podem ser vistas no diagrama de caso de uso do sistema representado pela Figura 5. Na Figura 6, é apresentada a organização da arquitetura e a comunicação dos seus componentes.

Figura 5: Diagrama de Caso de Uso UML

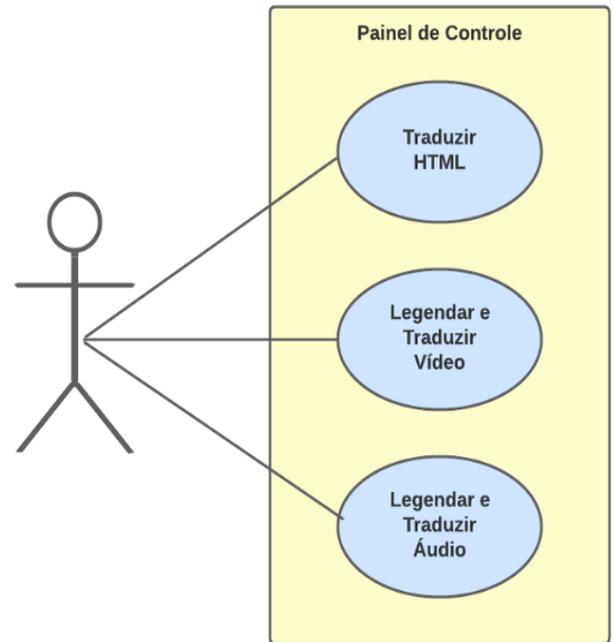
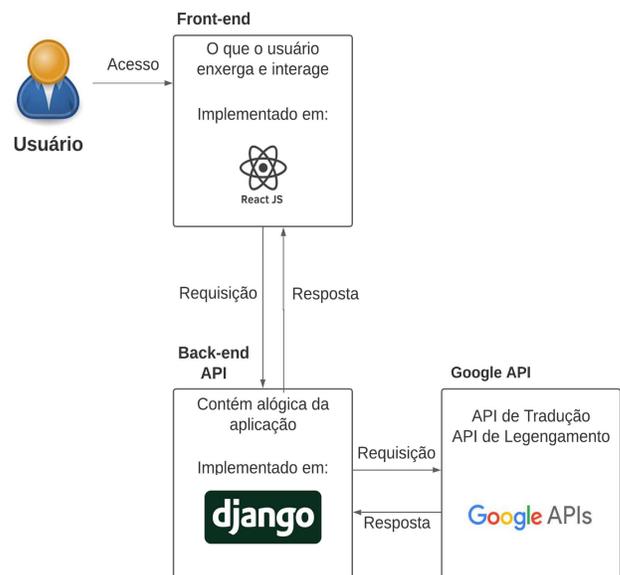


Figura 5: Arquitetura da Aplicação Web



Nas seções a seguir, serão descritos o processo de desenvolvimento adotado, as funcionalidades, a arquitetura utilizada e os resultados obtidos.

### 3.1 Processo de Desenvolvimento

No início do desenvolvimento do trabalho, foram realizadas reuniões com o orientador deste TCC, para compreender quais os principais problemas que a aplicação deveria resolver, levando em consideração o tempo de desenvolvimento e necessidades mais importantes. Após o processo de levantamento de requisitos e sua aprovação, as tecnologias a serem utilizadas foram definidas.

### 3.2 Arquitetura

A arquitetura de uma aplicação web é responsável por definir as interações entre aplicativos, *middlewares* e banco de dados. Os principais agentes da arquitetura utilizada são o *back-end* e o *front-end*, responsáveis respectivamente pelo servidor e cliente.

O servidor atua no recebimento, processamento e devolução de requisições, assim como manipulação do banco de dados, que por sua vez, tem como responsabilidade estruturar, criar, remover e editar dados.

O cliente é responsável por exibir uma interface visual para que o usuário possa interagir e utilizar as funcionalidades disponíveis na aplicação.

#### 3.2.1 Tecnologias do back-end

O servidor foi implementado utilizando o *framework Django* [6]. Essa tecnologia foi escolhida por diversos motivos, como por exemplo, o fato de ser código aberto, de permitir a integração de outras bibliotecas *Python* [7] na lógica da aplicação, assim possibilitando o processamento de dados vindo do *client*. Como também a facilidade de integração com o *React* [8].

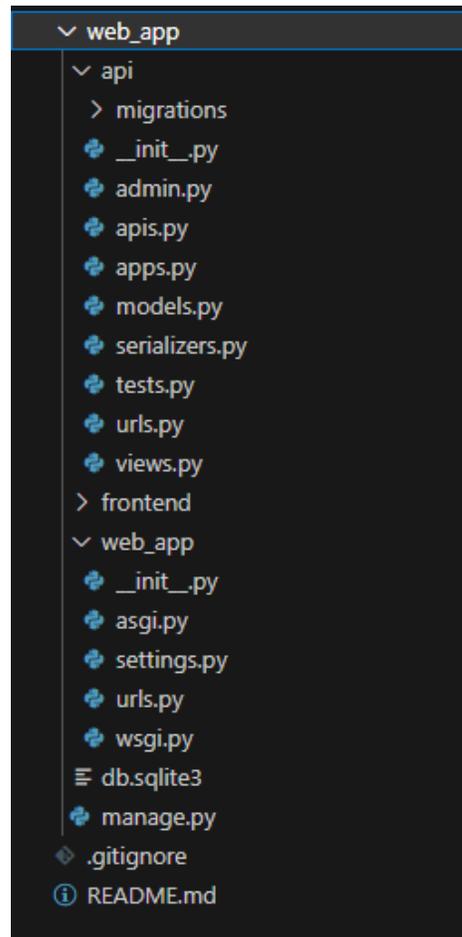
Para a tradução de texto em arquivos HTML, foi utilizada a biblioteca *py-googletrans-html* [10] (ou *Googletrans*), desenvolvida para a linguagem de programação *Python*. O *Googletrans* implementa a *API Google Translate* [11], utilizada para fazer chamadas para métodos de detecção de idiomas e tradução de textos.

Para a geração de legendas a partir de vídeo ou áudio, e a tradução das legendas geradas, foi utilizada a biblioteca *autosub* [12], desenvolvida para a linguagem de programação *Python*. O *Autosub* é um utilitário para reconhecimento automático de voz e geração de legendas. Ele usa um arquivo de vídeo ou áudio como entrada, executa a detecção de atividade de voz para encontrar regiões da fala, fazendo solicitações para o *Google Web Speech API* [13], assim gerando transcrições para essas regiões, e por fim, salvando o resultado em um arquivo de texto.

#### 3.2.2 Estrutura do back-end

O *back-end* da aplicação foi estruturado em duas pastas principais: *webapp* e *api*. Na Figura 6 é ilustrada a estruturação dos diretórios:

Figura 6: Estrutura de diretórios do *back-end*.



Em *webapp* estão os principais scripts que executam o *framework* da aplicação, desde o *deploy*, as URLs e os metadados. Também auxiliam na manipulação das rotas, controlando quais solicitações cada *controller* recebe para, em seguida, definir os *end-points* solicitados pelo cliente. Na pasta *api*, estão as bibliotecas para tradução de HTML, e transcrição de vídeo e áudio. Também há os scripts que controlam as requisições do *client*. Os dados podem ser transmitidos a partir do *body* das requisições, que podem ser do tipo POST, GET, PATCH e PUT e os parâmetros podem ser passados como *Path Parameter* ou *Query Parameter*.

#### 3.2.3 Tecnologias do front-end

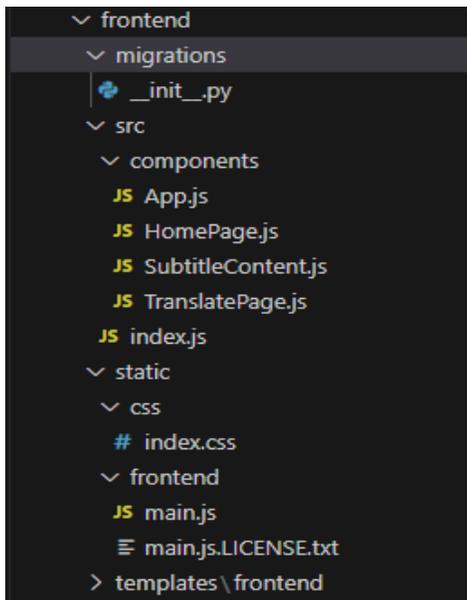
O *cliente* foi implementado utilizando *React* [8], uma biblioteca JavaScript de código aberto com foco na criação de interfaces de usuários em páginas web. Essa tecnologia foi escolhida devido à possibilidade de reutilização de componentes criados, evitando repetição de código. Além disso, a tecnologia fornece melhor modularização, permitindo que a aplicação seja mais bem organizada e escalável, ou seja, de fácil manutenção, independentemente do crescimento do seu tamanho e da complexidade.

Para tornar o processo de desenvolvimento mais rápido, utilizou-se a biblioteca de componentes *Material UI* [9], que já dispõe de componentes que oferecem consistência à interface do usuário e possibilita a estilização de layout responsivo de forma simples.

### 3.2.4 Estrutura do front-end

O *front-end* da aplicação foi dividido em três diretórios principais, sendo esses: *components*, *static* e *templates*. No diretório *components* encontram-se todas as páginas do sistema e seus respectivos componentes. Também encontram-se todos os componentes que são utilizados diversas vezes por outros componentes, como navbar, botões, campos de input, títulos, subtítulos, legendas e cards. Os diretórios *static* e *templates* contém as folhas de estilo, imagens e componentes que estruturam o HTML. Na Figura 7 é ilustrada a estruturação dos diretórios:

Figura 7: Estrutura de diretórios do front-end.



### 3.2.5 Páginas desenvolvidas

A seguir, a descrição das páginas desenvolvidas para o sistema web:

- **Página inicial:** Nesta página, representada pela Figura 8, há 3 botões para as seguintes funcionalidades: traduzir um arquivo *HTML*, legendar um arquivo de vídeo ou legendar um arquivo de áudio.
- **Página de tradução de HTML:** Na página inicial, ao clicar no botão “Traduzir HTML”, será encaminhado para a página de tradução de HTML, representada pela Figura 9. Nela há 3 botões, com as seguintes funcionalidades: ao clicar no botão “Definir Idioma”, será aberto um *pop-up*, como mostra a Figura 10, nela é possível definir o idioma da tradução; Ao clicar no botão “Carregar Arquivo”, poderá escolher o arquivo *HTML* que será enviado para o servidor; Por fim, ao clicar no botão “Traduzir HTML”, será enviado o

arquivo *HTML* para o servidor, o qual será processado, e fará o *download* automático do arquivo HTML traduzido.

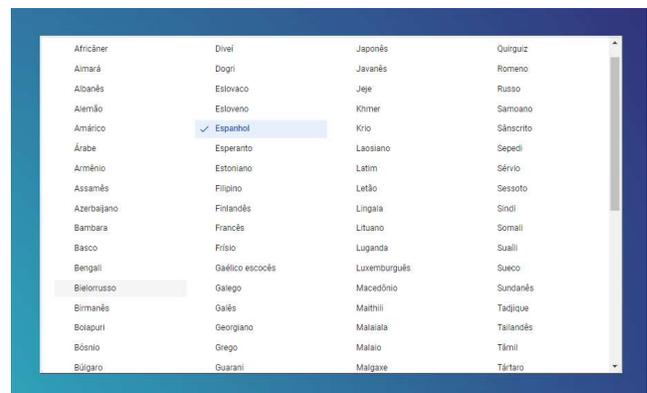
Figura 8: Página inicial do sistema web.



Figura 9: Página de tradução de HTML



Figura 10: Pop-up para definir o idioma de tradução



- **Página de transcrição de vídeo e tradução de legendas:** Na página inicial, ao clicar no botão “Legendar Vídeo”, será encaminhado para a página de tradução de HTML, representada pela Figura 11. Nela há 4 botões, com as seguintes funcionalidades: ao clicar no botão “Definir Idioma do Vídeo”, será aberto um *pop-up*, semelhante ao da Figura 10, onde é possível definir o idioma originário do vídeo que será transcrito; Ao clicar no botão “Definir Idioma da Tradução”,

também será aberto um *pop-up*, semelhante ao da Figura 10, nela é possível definir o idioma de tradução do vídeo transcrito, mas caso não queira traduzir, basta escolher o mesmo idioma originário do vídeo; Ao clicar no botão “Carregar Arquivo”, poderá escolher o arquivo de vídeo que será enviado para o servidor; Por fim, ao clicar no botão “Legendar e Traduzir Vídeo”, será enviado o arquivo de vídeo para o servidor, o qual será processado, e fará o *download* automático do arquivo que contém a transcrição e tradução das legendas.

Figura 11: Página para transcrever vídeo e traduzir legendas.



- Página de transcrição de áudio e tradução das legendas:** Na página inicial, ao clicar no botão “Legendar Áudio”, será encaminhado para a página de tradução de HTML, representada pela Figura 12. Nela há 4 botões, com as seguintes funcionalidades: ao clicar no botão “Definir Idioma do Áudio”, será aberto um *pop-up*, semelhante ao da Figura 10, onde é possível definir o idioma original do áudio que será transcrito; ao clicar no botão “Definir Idioma da Tradução”, também será aberto um *pop-up*, semelhante ao da Figura 10, nela é possível definir o idioma de tradução do áudio transcrito, mas caso não queira traduzir, basta escolher o mesmo idioma originário do áudio; Ao clicar no botão “Carregar Arquivo”, poderá escolher o arquivo de áudio que será enviado para o servidor; Por fim, ao clicar no botão “Legendar e Traduzir Vídeo”, será enviado o arquivo de áudio para o servidor, o qual será processado, e fará o *download* automático do arquivo que contém a transcrição e tradução das legendas.

Figura 12: Página para transcrever áudio e traduzir legendas.



## 4. EXPERIÊNCIA ADQUIRIDA

Desenvolver uma aplicação web completa foi uma experiência que proporcionou grande evolução das habilidades de programação e permitiu colocar em prática todo o conhecimento adquirido ao longo da graduação, em várias disciplinas cursadas, como por exemplo, análise de sistemas, engenharia de software, gestão de projeto, projeto de software, interface homem-máquina, entre tantas outras.

Cada etapa do desenvolvimento foi desafiadora, desde a elaboração dos requisitos até a pesquisa sobre o estado da arte da internacionalização de websites. Mesmo diante desses desafios, o projeto possibilitou a expansão do conhecimento acerca do processo de pesquisa e desenvolvimento de uma aplicação web.

## 5. CONSIDERAÇÕES FINAIS

Este trabalho teve como principal objetivo apresentar os principais desafios no desenvolvimento de websites, considerando os princípios da internacionalização e localização de websites. Assim como, o desenvolvimento de uma aplicação web capaz de auxiliar na tradução e legendamento do conteúdo de um website, focando no HTML, vídeo e áudio.

Diante do exposto, o trabalho desenvolvido, proporcionou não apenas o fortalecimento de conhecimentos na área de desenvolvimento web, como também o desenvolvimento nas habilidades de pesquisa.

## 6. REFERÊNCIAS

- [1] Website. Wikipedia, 2023. Disponível em: [https://en.wikipedia.org/wiki/Website#:~:text=A%20website%20\(also%20written%20as,commerce%2C%20entertainment%20or%20social%20networking](https://en.wikipedia.org/wiki/Website#:~:text=A%20website%20(also%20written%20as,commerce%2C%20entertainment%20or%20social%20networking). Acesso em: 25 de maio de 2023.
- [2] FRY, Deborah. Cartilha da Indústria de Localização. 2ª edição. Lisa, n. 2, v. 5, p. 8, 2003. Disponível em: <https://www.immagic.com/eLibrary/ARCHIVES/GENERAL/LISA/L030625P.pdf>.
- [3] ESSELINK, Bert. Um Guia Prático para Localização de Software. Amsterdã: John Benjamins, n. 3, v. 12, p. 35, 2001. Disponível em: <https://benjamins.com/catalog/liwd.4>.
- [4] YUNKER, John. Além das Fronteiras: Estratégias de Globalização. Indianapolis: New Riders, n. 5, v. 3, p. 42, 2003. Disponível em [https://books.google.com.br/books/about/Beyond\\_Borders.html?id=tJ0JcoLfeloC&redir\\_esc=y](https://books.google.com.br/books/about/Beyond_Borders.html?id=tJ0JcoLfeloC&redir_esc=y).
- [5] Hospedagem de sites web. Wikipedia, 2022. Disponível em [https://pt.wikipedia.org/wiki/Hospedagem\\_de\\_s%C3%ADtio\\_s\\_web](https://pt.wikipedia.org/wiki/Hospedagem_de_s%C3%ADtio_s_web). Acesso em: 10 de jun. de 2023.
- [6] Django Framework: tecnologia utilizada no *back-end*. Disponível em: <https://www.djangoproject.com/>. Acesso em: 30 de maio de 2023.
- [7] Python: tecnologia utilizada no *back-end*. Disponível em: <https://www.python.org/>. Acesso em: 12 de jun. de 2023.
- [8] React: tecnologia utilizada no *front-end*. Disponível em: <https://react.dev/>. Acesso em: 12 de jun. de 2023.

- [9] Material UI: tecnologia utilizada no desenvolvimento do *front-end*. Disponível em: <https://mui.com/>. Acesso em: 12 de jun. de 2023.
- [10] Googletrans: biblioteca utilizada para traduzir texto presente em um arquivo HTML. Disponível em: <https://pypi.org/project/py-googletrans-html/>. Acesso em: 12 de jun. de 2023.
- [11] API Google Translate: utilizada para traduzir textos. Disponível em: <https://cloud.google.com/translate/?hl=pt-br>. Acesso em: 12 de jun. de 2023.
- [12] Autosub: utilizado para a geração de legendas de vídeos e áudios. Disponível em: <https://pypi.org/project/autosub/>. Acesso em: 12 de jun. de 2023.
- [13] Google Web Speech API: serviço da Google para reconhecimento de voz e geração de transcrição de áudio ou vídeo. Disponível em: <https://cloud.google.com/speech-to-text?hl=pt-br>. Acesso em: 12 de jun. de 2023.
- [14] RESENDE, Helena; Silva, Manuel. Gestão de projetos de tradução e de localização. Instituto Politécnico do Porto. Instituto Superior de Contabilidade e Administração do Porto, 2009. Disponível em: <https://recipp.ipp.pt/handle/10400.22/2883>.
- [15] ALVES, Fernando Ferreira. Da profissão à formação: o enquadramento da formação de tradutores em contexto de trabalho. Lisboa, Portugal, n. 1, v. 2, p. 11, 2005. Disponível em: <https://repositorium.sdum.uminho.pt/handle/1822/7175>.
- [16] DARIN, Leila Cristina de Mello. A tradução no contexto do mundo globalizado. PUC, São Paulo, n. 2, v. 5, p. 8, 2010. Disponível em <https://revistas.pucsp.br/intercambio/article/view/3959>.
- [17] Adobe PDF: sobre arquivos de PDF. Disponível em: <https://www.adobe.com/br/acrobat/pdf-reader.html>. Acesso em: 12 de jun. de 2023.