

CURSO DE SISTEMAS DE INFORMAÇÃO – TCC ACADÊMICO	
() PRÉ-PROJETO (X) PROJETO	ANO/SEMESTRE: 2020/2

YOUBORDERLESS: UNINDO LINGUAGENS E CULTURAS

Hugo Marcel Larsen

Prof.^a Simone Erbs da Costa – Orientadora

1 INTRODUÇÃO

Em 11 de março de 2020, a Organização Mundial da Saúde (OMS) declarou que o novo Corona Vírus (nCoV-2019 ou Sars-cov-2), causador da doença COrona Vírus Disease 2019 (COVID-19), se tornou uma pandemia (BARROZO *et al.* 2020). Para monitorar o surto atual, medidas robustas foram tomadas em todo o mundo para reduzir a transmissão do COVID-19, especialmente proibindo voos internacionais, criando protocolos para *lockdowns* em áreas vulneráveis e aplicando o distanciamento social. (HAMID *et al.*, 2020).

Segundo Deslandes e Coutinho (2020), durante o isolamento social, o acesso à internet foi crescendo exponencialmente, sendo o principal meio de interatividade entre as pessoas. Com base na diminuição social entre as pessoas e suas drásticas mudanças de cotidiano, espera-se que o uso da internet sirva de normalizador de situações que antes não eram possíveis off-line e não agravem os efeitos colaterais que o isolamento social pode causar, como: depressão, ansiedade, solidão, violências familiares e suicídios (DESLANDES; COUTINHO, 2020).

Nesse sentido, o Youtube é amplamente utilizado como suporte pedagógico por uma quantidade significativa de internautas brasileiros conforme pesquisas, sendo a plataforma mais acessada para fins de conhecimento (18%), seguida da TV Globo (17%) e do Facebook (8%) (NAGUMO; TELES; SILVA, 2020). Xavier *et al.* (2019) também colocam o quanto o aparecimento de mídias e tecnologias, como o Youtube, ajudam tanto pessoas ouvintes como surdas, superando assim limites impostos no processo tradicional de ensino. Xavier *et al.* (2019) ainda destacam a importância do uso do Youtube neste momento, devido a possibilidade de professores e educadores o usarem visando uma maior inclusão social quando comparadas as ferramentas tradicionais de educação.

Diante deste cenário, este trabalho propõe o desenvolvimento de um protótipo que apoiará pessoas a entenderem transmissões ao vivo no Youtube, com auxílio de legendas. Conjectura-se que a construção deste protótipo quebre a barreira linguística entre conteúdos de interesse pessoal e auxilie em engajamentos interculturais e estudantis. Assim, objetivando também amenizar os efeitos colaterais citados causados durante o período da pandemia do COVID-19.

1.1 OBJETIVOS

O objetivo geral do trabalho é disponibilizar um protótipo para auxiliar pessoas a entenderem transmissões ao vivo pelo Youtube com a ajuda de legendas em seu idioma. Sendo os objetivos específicos:

- a) disponibilizar uma aplicação em ambiente web e móvel;
- b) promover a aprendizagem por meio de transcrições de transmissão ao vivo pelo Youtube, assim como possibilitar a interação entre diferentes culturas, amenizando os efeitos colaterais causados durante o período da pandemia do COVID-19;
- c) modelar a relação entre os requisitos da aplicação e práticas consolidadas no design de interface como as heurísticas de Nielsen, pelo método Relationship of M3C with User Requirements and Usability and Communicability Assessment in groupware (RURUCAg), possibilitando avaliar de maneira simples a usabilidade das aplicações disponibilizadas.

2 TRABALHOS CORRELATOS

Nesta seção serão descritos três projetos correlatos que apresentam características diretamente relacionadas ao trabalho proposto. A subseção 2.1 detalha o uso de um microcontrolador para transcrever a voz humana por meio de um hardware criado e um dispositivo Android (UKO; DUBUKUMAH; AYOSUBOMI, 2019). A subseção 2.2 apresenta o desenvolvimento de uma aplicação de computador para traduzir e transcrever o idioma filipino para o idioma chinês (BAILON *et al.*, 2019). Por fim, a subseção 2.3 descreve a aplicação de computador KT-Speech-Crawler, na qual obtém dados de vídeos do Youtube para treinamento de Redes Neurais Artificiais (RNA) (LAKOMKIN *et al.*, 2018).

2.1 MICROCONTROLLER BASED SPEECH TO TEXT TRANSLATION SYSTEM

Uko, Dubukumah e Ayosubomi (2019) propuseram uma solução para resolver os problemas de uma comunicação efetiva em tempo real entre pessoas. No desenvolvimento do projeto foi utilizado o microcontrolador ATMEGA328p como recurso principal e “cérebro” do projeto; uma aplicação Android para captar a voz por um dispositivo móvel; um visor Liquid Crystal Display (LCD) para exibir o texto traduzido como a saída final para o usuário; um módulo Bluetooth para transmitir os dados da voz da aplicação Android; e outros componentes como provedores de energia (UKO; DUBUKUMAH; AYOSUBOMI, 2019).

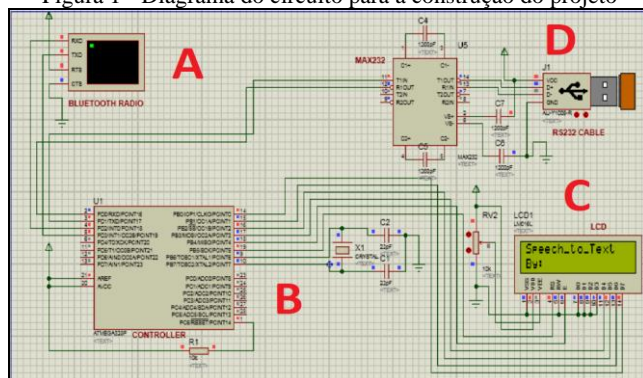
Segundo Uko, Dubukumah e Ayosubomi (2019), o trabalho teve como objetivo transcrever a fala humana em texto com um projeto de baixo custo. O intuito foi apresentar a

metodologia, limitações, recomendações e aplicações para melhorar a comunicação humana com a integração de dispositivos eletrônicos, no qual não foi bem explorada devido complexidades como a captação de variações de sons para reconhecimento (UKO; DUBUKUMAH; AYOSUBOMI, 2019). Cabe destacar, algumas das características de Uko, Dubukumah e Ayosubomi (2019), como: reconhecimento de voz, transcrição de voz e disponibilização tanto na plataforma Android como por meio do hardware criado.

O processo da transcrição de voz se inicia com a entrada da fala humana por meio do microfone no dispositivo Android. Duas aplicações são necessárias no dispositivo: o Adaptive Multi-Rate (AMR), um esquema de compressão de áudio usado para a codificação da voz; e o Google Voice Search utilizado como suporte para obter o vocabulário da linguagem a ser transcrita. Assim, os caracteres da linguagem reconhecida pelas aplicações são transferidos por comunicação serial por meio do módulo Bluetooth para o microcontrolador. Os dados dos caracteres precisam por fim serem convertidos de analógicos para digitais para serem exibidos nos visores (UKO; DUBUKUMAH; AYOSUBOMI, 2019).

O diagrama do circuito do projeto (Figura 1) demonstra como o aparelho desenvolvido tem seus módulos organizados. Na letra A pode ser visto o módulo de rádio Bluetooth que recebe as informações transferidas do dispositivo Android. A partir daí, o microcontrolador ATMEGA328p (letra B), “cérebro” do projeto, processa todas as funções programadas no ambiente Arduino e compartilha suas informações com os outros módulos. Após cálculos e funções executadas, as letras C e D exibem o texto transcrito ao usuário, ou seja, a informação final. A letra C destaca o visor LCD incorporado no próprio dispositivo e a letra D a utilização do cabo RS 232 para exibição em visores maiores por exemplo (UKO; DUBUKUMAH; AYOSUBOMI, 2019).

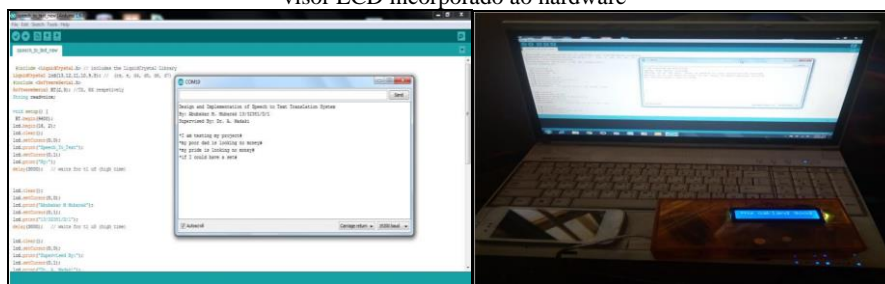
Figura 1 - Diagrama do circuito para a construção do projeto



Fonte: adaptada de Uko, Dubukumah e Ayosubomi (2019).

Uko, Dubukumah e Ayosubomi (2019) observam as limitações quanto ao uso do dispositivo desenvolvido, como: alcance da área de Bluetooth, tanto do dispositivo Android quanto do módulo Bluetooth; ruídos e barulhos que podem atrapalhar a captação genuína da voz; falta de mais telas para armazenamento do texto que já foi exibido transcrito; somente falas com um bom sotaque inglês eram traduzidas para texto; e o projeto foi limitado apenas com a língua inglesa. Ainda é possível observar a transcrição da voz humana sendo exibida de duas formas: no ambiente Arduino (Figura 2 (a)); e no próprio visor LCD incorporado no hardware (Figura 2 (b)) (UKO; DUBUKUMAH; AYOSUBOMI, 2019).

Figura 2 – (a) Voz humana transcrita no ambiente Arduino e (b) Voz humana transcrita no visor LCD incorporado ao hardware



Fonte: adaptada de Uko, Dubukumah e Ayosubomi (2019).

Uko, Dubukumah e Ayosubomi (2019) colocam que por meio da realização desse projeto novas possibilidades e avanços na pesquisa de transcrição de texto são abordadas, como: fornecimento de um método renovado para ser aplicado na comunicação de dispositivos eletrônicos, como caixas eletrônicas, dispositivos multimídia ou robôs; aplicação em automação residencial e aplicações de segurança.

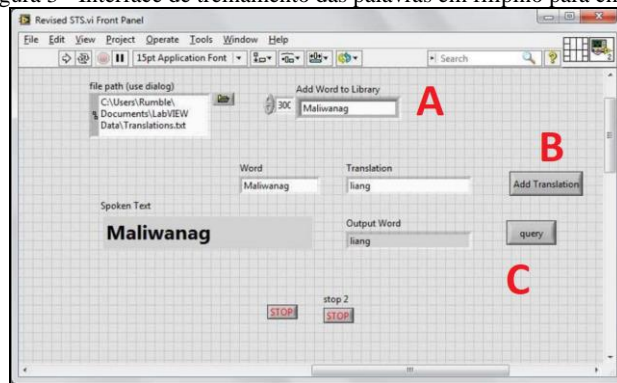
2.2 FILIPINO TO CHINESE SPEECH-TO-SPEECH TRANSLATOR USING NEURAL NETWORK WITH DATABASE SYSTEM

Bailon *et al.* (2019) tiveram como objetivo desenvolver uma aplicação que usa um algoritmo para traduzir a fala do idioma filipino para a fala no idioma chinês. O controlador principal da aplicação é uma Inteligência Artificial (IA) baseada em uma RNA, auxiliada pela LabVIEW. As três etapas principais da aplicação são: reconhecimento da fala filipina para transcrição; tradução de texto filipino para chinês; reprodução falada do texto chinês (BAILON *et al.*, 2019). Cabe destacar, algumas das características de Bailon *et al.* (2019) como: reconhecimento de voz, transcrição de voz, tradução de texto e disponibilização para a plataforma desktop.

O fluxo do processo de estudo é separado em três estágios, sendo eles: de gravação; de recolhimento; e de testagem. No estágio de gravação são armazenadas 25 variações de fala de uma palavra. Isso é feito para 500 palavras frequentemente usadas no idioma filipino. Para o estágio de coleta, 500 palavras em filipino terão de corresponder a 500 palavras em chinês. Esse estágio também pode ser chamado de treinamento porque as variações de fala serão armazenadas na RNA e ela vai ser treinada para que reconheça as palavras específicas, independente do padrão de fala e ruídos externos (BAILON *et al.*, 2019).

Para ajudar no treinamento da RNA, a aplicação e a interface foram desenvolvidas para terem as seguintes funcionalidades (Figura 3): adicionar a palavra ao dicionário (letra A); inserir a tradução da palavra em filipino para chinês (letra B); utilizar um microfone para gravar a fala do usuário para a palavra correspondente (letra C). Nessa etapa, o papel do Windows Speech Recognition foi fundamental para deixar o reconhecimento da voz mais preciso, assim aumentando a acurácia em relação ao treinamento das palavras armazenadas na RNA (BAILON *et al.*, 2019).

Figura 3 - Interface de treinamento das palavras em filipino para chinês



Fonte: adaptada de Bailon *et al.* (2019).

Os resultados apresentados por Bailon *et al.* (2019) mostram lacunas em algumas palavras que não obtiveram uma taxa de sucesso alta na tradução porque possuem pronúncias parecidas. Entretanto, Bailon *et al.* (2019) observaram que a maioria das palavras teve baixo índice de rejeição, algumas possuindo a pontuação perfeita. Bailon *et al.* (2019) ainda colocam que para melhorar a taxa de sucesso das traduções, principalmente nas palavras com pontuação baixa, treinamentos com inúmeras iterações foram realizados. Com isso, a taxa de sucesso, durante a etapa de treinamento, foi aumentada significativamente conforme os dados da Tabela 1. Entretanto, durante a fase de teste, realizada com 25 pessoas, a aplicação teve uma taxa de sucesso de 91,59% com uma taxa de rejeição de 8,41% (BAILON *et al.*, 2019).

Tabela 1 - Taxa de sucesso das palavras

Número de iterações	Percentual de sucesso
25	86.8
37	88.2
49	89.6
61	91.8
73	92.0
85	93.4

Fonte: adaptada de Bailon *et al.* (2019).

Bailon *et al.* (2019) trazem recomendações propostas pelos pesquisadores da pesquisa, sendo elas: mais variações de voz por palavra, que conforme comprovado, aumentam a eficácia de reconhecimento de palavras, juntamente com o número de iterações de treinamento; aumentar a eficiência do reconhecimento da palavra de entrada para torná-la em torno de 97 a 99 por cento e um por cento destinado a variáveis externas incontroláveis; aumento de duas a três mil palavras, englobando mais do vocabulário, gírias e expressões; e, a longo prazo, uma aplicação expandida da RNA para reconhecimento de frases e sentenças, chegando assim a uma gramática correta e mais próximo do objetivo de ter um tradutor totalmente controlado por computador (BAILON *et al.*, 2019).

2.3 KT-SPEECH-CRAWLER: AUTOMATIC DATASET CONSTRUCTION FOR SPEECH RECOGNITION FROM YOUTUBE VIDEOS

Para melhorar aplicações de reconhecimento de voz, Lakomkin *et al.* (2018) criaram o KT-Speech-Crawler, desenvolvido na linguagem Python, que tem o objetivo de recuperar informações de vídeos do Youtube. Essas informações são usadas para treinamento de RNA. A YouTube Search Application Programming Interface (API) foi a principal tecnologia utilizada para recuperar os vídeos do Youtube e suas legendas ocultas (LAKOMKIN *et al.*, 2018). Cabe destacar, algumas das características de Lakomkin *et al.* (2018), como: recuperar informações, reconhecimento de voz, transcrição de voz e disponibilização para a plataforma desktop.

O *crawler* pode obter cerca de 150 horas de fala transcrita em um dia, contendo uma taxa de erro de 3,5% no reconhecimento. As falas coletadas são gravadas em diversas condições, incluindo ruídos, músicas, má qualidade da voz por estar longe do microfone e diferentes sotaques e entonações. Foi observado uma redução de 40% na taxa de erros de palavras do conjunto de dados do Wall Street Journal apenas integrando 200 horas de amostras coletadas pelos vídeos no conjunto de treinamento (LAKOMKIN *et al.*, 2018).

As etapas principais executadas pelo KT-Speech-Crawler estão demonstradas na Figura 4 (a): processo de pesquisa dos vídeos pela API do Youtube, com palavras chaves da língua

inglesa; na letra B são removidos, por exemplo, textos indesejados e músicas de fundo para os filtros de aplicação, tanto no áudio quanto nas legendas ocultas; na etapa de pós processamento (letra C) o programa tenta corrigir desalinhamentos imprecisos entre o áudio e as legendas do vídeo; a penúltima etapa (letra D) traz a transcrição do vídeo; e por fim, na letra E, a transcrição é enviada para a RNA ser treinada (LAKOMKIN *et al.*, 2018). Já a Figura 4 (b) traz a interface web desenvolvida por Lakomkin *et al.* (2018) para analisar as amostras retiradas e suas transcrições. O *web service* desenvolvido exibe oito exemplos aleatórios e sua correspondente transcrição, permitindo carregar mais se necessário. Também foi integrada uma funcionalidade para que o usuário confirme se a transcrição está correta ou se não corresponde ao que foi falado no áudio. Lakomkin *et al.* (2018) estimam uma taxa de erro de 3,5% para um total de 600 amostras aleatórias.

Figura 4 – (a) Etapas do processo executado pelo KT-Speech-Crawler e (b) Interface web para análise das amostras e suas transcrições



Fonte: adaptada de Lakomkin *et al.* (2018).

Na demonstração de resultados, Lakomkin *et al.* (2018) provaram que adicionando amostras capturadas do Youtube aumentaram positivamente a performance do reconhecimento de voz dos dois conjuntos de dados, Wall Street Journal e TED-LIUM v2. Foram adicionadas 200 horas de amostras do Youtube (108.617 dicções) no conjunto de treinamento do Wall Street Journal e a taxa de erro de palavras foi melhorada de 34,2% para 15,8%. Um resultado similar foi visto no conjunto de dados TED-LIUM v2 ao adicionar 300 horas de amostras do Youtube, melhorando a taxa de erro de caracteres de 10,4% para 8,2% (LAKOMKIN *et al.*, 2018).

3 PROPOSTA DO PROTÓTIPO

Nesta seção serão apresentadas as justificativas para a realização do trabalho proposto (subseção 3.1). Na subseção 3.2 serão descritos os requisitos principais. Finalizando com a

subseção 3.3, serão descritos o cronograma e a metodologia planejados para o desenvolvimento do trabalho.

3.1 JUSTIFICATIVA

Nas seções 1 e 2 foram evidenciadas a relevância do tema proposto. No Quadro 1 é apresentado um comparativo entre os trabalhos correlatos, de modo que as linhas representam as características e as colunas os trabalhos relacionados.

Quadro 1 - Comparativo entre os correlatos

Características	Correlatos	Uko, Dubukumah e Ayosubomi (2019)	Bailon <i>et al.</i> (2019)	Lakomkin <i>et al.</i> (2018)
Recuperar informações de vídeos do Youtube		X	X	✓
Reconhecimento de voz		✓	✓	✓
Transcrição de voz		✓	✓	✓
Tradução de texto		X	✓	X
Plataforma		Hardware	Desktop	Desktop

Fonte: elaborado pelo autor.

Conforme demonstrado no Quadro 1, percebe-se que Uko, Dubukumah e Ayosubomi (2019), Bailon *et al.* (2019) e Lakomkin *et al.* (2018) têm finalidades semelhantes no que se diz respeito a reconhecimento e transcrição de voz. Ambas as características foram identificadas nos três correlatos como essenciais para alcançar seus objetivos, uma vez que são importantes para a operacionalização das demais características. Lakomkin *et al.* (2018) se destacam por apresentarem a característica relacionada a recuperar informações de vídeos do Youtube, dizendo respeito a recuperar dados da voz e outras informações.

O reconhecimento de voz é uma característica que foi vista em todos os trabalhos correlatos. Em Uko, Dubukumah e Ayosubomi (2019) a voz era recuperada por um dispositivo Android e transmitida pelo módulo Bluetooth para o hardware criado. Para treinar a RNA, Bailon *et al.* (2019) utilizaram o Windows Speech Recognition a fim de reconhecer e melhorar o reconhecimento de voz por parte dos participantes. Já Lakomkin *et al.* (2018) fizeram uso das informações de vídeos do Youtube para obter a voz e treinar RNA.

Assim como a característica de reconhecimento de voz, a transcrição de voz também é fundamental nos três correlatos. Uko, Dubukumah e Ayosubomi (2019) precisam dessa característica para que o usuário saiba o que foi dito por outra pessoa. Bailon *et al.* (2019) necessitam transcrever a voz filipina para que seja possível traduzir o texto para chinês e manter a comunicação desses dois idiomas. Lakomkin *et al.* (2018) fizeram uso dessa característica para treinar RNA e conferir os resultados obtidos das amostras.

A característica de tradução de texto foi vista somente em Bailon *et al.* (2019). Ela se assemelha a este protótipo porque a pessoas que não falam o mesmo idioma possam entender

o conteúdo da transmissão ao vivo. Bailon *et al.* (2019) têm seu principal objetivo voltado para essa característica, na qual a comunicação entre filipino e chinês possa ser compreendida com resultados satisfatórios.

Com base nestas características e no Quadro 1 apresentado, nota-se que o trabalho proposto se torna relevante ao buscar ajudar pessoas a entenderem transmissões ao vivo que não são de seu idioma. Além disso, a proposta visa auxiliar em outros aspectos como a integração de novos conteúdos na vida das pessoas para que possam aprofundar estudos e até mesmo se divertirem ao entenderem o que está acontecendo em transmissões ao vivo no Youtube. Já como contribuição acadêmica, serão criadas interfaces amigáveis para serem aplicadas ao protótipo e validadas com a utilização do Método RURUCAg. Deverá ser avaliada a experiência e usabilidade de usuários na aplicação e relacionados guias consolidados de criação de interfaces como as heurísticas de Nielsen. Como contribuição tecnológica pode-se destacar o desenvolvimento de uma aplicação web e uma aplicação móvel que facilitarão o entendimento de vídeos ao vivo no Youtube com auxílio de legendas. Serão utilizadas tecnologias em nuvem para transcrição e tradução de informações, juntamente com a tecnologia Progressive Web Apps (PWA) para flexibilização e escalabilidade, além de fazer uso de um Sistema Gerenciador de Banco de Dados (SGBD) Not Only Structured Query Language (NoSQL).

3.2 REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

Nesta subseção serão abordados os principais Requisitos Funcionais (RF), assim como os principais Requisitos Não Funcionais (RNF), conforme Quadro 2.

Quadro 2 – Principais Requisitos Funcionais e Não Funcionais

A aplicação deverá:	Tipo
permitir ao usuário acesso por meio de uma conta Google	RNF
permitir ao usuário consultar uma lista de vídeos ao vivo do Youtube conforme as inscrições do usuário autenticado com a conta Google	RF
permitir ao usuário visualizar um vídeo ao vivo	RF
permitir ao usuário configurar a transcrição para o vídeo ao vivo sendo visualizado	RF
utilizar serviços Google Cloud para transcrever áudio e traduzir essa transcrição	RNF
utilizar APIs para recuperar informações de vídeos ao vivo do Youtube	RNF
implementar a aplicação web utilizando o <i>framework</i> Angular	RNF
implementar a aplicação web seguindo as diretrizes do Material Design	RNF
consultar serviços na parte do servidor desenvolvido em NodeJS	RNF
implementar uma aplicação móvel com a tecnologia PWA	RNF
implementar uma aplicação móvel com a ajuda do ambiente de desenvolvimento Visual Studio Code	RNF
utilizar a plataforma Firebase (Google) para usar serviços de hospedagem e validar acessos de contas do Google	RNF
utilizar a plataforma Firebase como SGBD NoSQL para armazenar informações dos usuários	RNF
utilizar o Método RURUCAg para modelar os requisitos da aplicação com as heurísticas de Nielsen	RNF
utilizar o Método RURUCAg para avaliar a usabilidade e a experiência de uso da aplicação	RNF

Fonte: elaborado pelo autor.

3.3 METODOLOGIA

A metodologia desta proposta está elaborada em seis etapas e composta pelos instrumentos metodológicos:

- pesquisa na literatura: realizar uma revisão mais aprofundada da literatura sobre os assuntos citados na revisão bibliográfica e trabalhos correlatos;
- refinamento dos requisitos: reavaliar os requisitos funcionais e não funcionais já definidos e, se necessário;
- especificação e análise: formalizar as funcionalidades da ferramenta por meio de casos de uso e diagramas de atividade da Unified Modeling Language (UML), utilizando a ferramenta Astah Community;
- implementação da aplicação web: implementar uma aplicação web para os usuários finais utilizando como base a linguagem de programação JavaScript e Visual Studio Code como ambiente de desenvolvimento;
- implementação da aplicação móvel: implementar uma aplicação móvel com a tecnologia PWA para exibir o conteúdo da aplicação web e Visual Studio Code como ambiente de desenvolvimento;
- verificação e validação: validar a usabilidade da solução pelo Método RURUCAg.

As etapas serão realizadas nos períodos relacionados no Quadro 3.

Quadro 3 – Cronograma

Etapas	Quinzenas		2021									
			fev.		mar.		abr.		maio		jun.	
	1	2	1	2	1	2	1	2	1	2	1	2
Pesquisa na literatura												
Refinamento dos requisitos												
Especificação e análise												
Implementação da aplicação web												
Implementação da aplicação móvel												
Verificação e validação												

Fonte: elaborado pelo autor.

4 REVISÃO BIBLIOGRÁFICA

Nesta seção são apresentados os conceitos e fundamentos mais importantes para a pesquisa em questão. A subseção 4.1 contextualiza a utilização do Youtube e recuperação de suas informações. A subseção 4.2 aborda temas de reconhecimento e tradução de fala. A subseção 4.3 contextualiza o conceito de usabilidade. Por fim, a subseção 4.4 descreve a metodologia de desenvolvimento PWA.

4.1 YOUTUBE E APRENDIZADO

Segundo Bardakci (2019), o Youtube é uma das mídias sociais mais prevalentes e um dos recursos digitais mais comuns na práxis educacional no mundo todo. Conforme os dados da Pesquisa Nacional por Amostra de Domicílios Contínua (PNADC) realizada no ano de 2016, uma das principais atividades do brasileiro na internet é assistir vídeos on-line no celular (IBGE, 2018). Segundo Nagumo, Teles e Silva (2020), a pesquisa Video Viewers realizada com três mil pessoas pelo Instituto Provokers, em 2018, mostrou que nove em cada dez brasileiros usam o Youtube para aprendizado.

Os resultados comprovados da pesquisa de Bardakci (2019), sobre a utilização do Youtube como aprendizado por estudantes do ensino médio, mostram que o fator mais significativo para tal não foram as condições facilitadoras e sim a intenção comportamental juntamente com a expectativa de desempenho, contrariando o que foi previsto no estudo. Similar ao que foi comprovado por Bardakci (2019), a pesquisa de Khechine e Lakhal (2018) mostrou que a intenção comportamental e a expectativa de desempenho também são fatores significantes na utilização da tecnologia para aprendizado por estudantes universitários. Foi investigado a aceitação da tecnologia de *webinar* por estudantes universitários e os resultados do estudo demonstraram que as condições facilitadoras não foram um fator significativo (KHECHINE; LAKHAL, 2018).

Conforme Malik e Tian (2017), o Youtube é o site de compartilhamento de vídeo mais popular e bem sucedido, os vídeos da plataforma se tornaram um tesouro de informações que podem ser usados em vários campos de pesquisa, desde educação Science, Technology, Engineering and Mathematics (STEM) a ciências médicas. A Youtube Data API (YDA) permite que os usuários realizem diversas operações como: procurar vídeos; visualizar canais; fazer upload de seus vídeos fora do Youtube; e até personalizar uma aplicação com base nas informações de vídeos do usuário (ADAM; AIMAAN SULAIMAN; SOH, 2019). Com a ajuda da YDA, que é fornecida pelo Youtube para desenvolvedores, durante um período de dois meses aplicando a metodologia criada, foram descobertos 16 milhões de vídeos e extraídas informações completas de mais de 42 mil vídeos (MALIK; TIAN, 2017).

4.2 RECONHECIMENTO E TRADUÇÃO DE FALA

O Google Tradutor é possivelmente o serviço de tradução mais utilizado e em 2016, mais de 500 milhões de pessoas estavam traduzindo 100 bilhões de palavras por dia em 103 idiomas, sendo o Brasil com maior número de usuários (BROOKS, 2016). O Google AssistantAPI é um serviço de assistente virtual, criado pela Google, para ajudar os usuários

em suas atividades diárias utilizando uma RNA para fazer perguntas, recomendações e executar comandos enviando solicitações a serviços web (ALBERTUS; MULIADY, 2020).

Albertus e Muliady (2020) percebem a importância do reconhecimento de fala para enviar comandos a um ar-condicionado por pessoas com deficiência que não conseguem operar um controle remoto de forma física. O reconhecimento de fala é amplamente utilizado para ajudar pessoas com deficiências, como por exemplo, controlar interruptores de luz ou até mesmo controlar movimentos de uma cadeira de rodas (ALBERTUS; MULIADY, 2020). Para automatizar mais esses processos e aproveitar a conveniência da transmissão de informação entre dispositivos eletrônicos, existem diferentes tecnologias que podem auxiliar a implementação de fala e uma delas é a de Sistemas de Redes Neurais Artificiais (SRNA) (BAILON *et al.*, 2019).

Ao se tratar do reconhecimento de fala, também é possível abordar a tradução de fala. Como uma constante no desenvolvimento da humanidade, a tradução sempre desempenhou um papel crucial na comunicação de diversos idiomas, permitindo o compartilhamento de conhecimento e cultura entre diferentes línguas (DOHERTY, 2016). Silva e Fernandes (2020), por exemplo, desenvolveram uma aplicação com base em um corpus paralelo para aperfeiçoar a pesquisa, ensino e prática da tradução. A aplicação permite a comparação da tradução automática, fornecida por serviços amplamente globalizados como Google Translate e Microsoft Translator, com a tradução humana captada pela aplicação para tradução dos corpus (SILVA; FERNANDES, 2020). Doherty (2016) ainda mostra que, ao explorar o impacto das tecnologias de tradução na comunicação internacional, de uma perspectiva interacionista, os efeitos sobre o processo de tradução, seus resultados e seu lugar na sociedade são notavelmente efetivos.

4.3 USABILIDADE E EXPERIÊNCIA DO USUÁRIO

Para se ter uma maior interação e feedback com o usuário, o uso de interfaces amigáveis é indispensável em uma aplicação desenvolvida, independente da tecnologia utilizada (COSTA, 2018). Ao encontro dessa vertente, está o Material Design (MD) criado pela Google para melhorar a experiência dos usuários. De acordo com Google (2019), o MD é um sistema de design criado para ajudar equipes a criar experiências digitais de alta qualidade para aplicações móveis e web.

O MD é um sistema adaptável de diretrizes, componentes e ferramentas que oferecem suporte às melhores práticas de design de interface do usuário, ajudando equipes a criarem produtos bonitos rapidamente (GOOGLE, 2019). O MD foi criado para ser referência em

interação com usuário e os desenvolvedores que procuram adequar a identidade visual de seus produtos em diferentes plataformas (TORRES *et al.*, 2018). Torres *et al.* (2018) ainda ressaltam a utilização do MD em seu trabalho como guia específico para acessibilidade, utilizando boas práticas de design voltadas para o público com deficiência. Adotando princípios da vida real, o MD oferece recursos ricos em gestos naturais que tendem a imitar objetos do mundo real, concentrando esse design em aplicativos móveis baseados em toque, embora seja possível interpretar a mesma ideia no design web (AZKYA, ARDIANSAH; e PUJANTO, 2020).

No processo de desenvolvimento de interfaces digitais, a experiência e usabilidade do usuário de sistemas de informação são de grande relevância (MAIA; BARBOSA; WILLIAMS, 2020). De acordo com Maia, Barbosa e Williams (2020), o conceito de experiência do usuário tem uma maior amplitude quando relacionado à usabilidade, na qual está fortemente ligada aos fatores maioritariamente instrumentais, relacionados à eficiência da tarefa ou do trabalho. Para avaliar a interface de aplicações e as suas funcionalidades, com base nas heurísticas de Nielsen, e assim reconhecer incidentes com relação a usabilidade e o uso da aplicação, foi criado o Método RURUCAg (COSTA, 2018). De acordo com Costa (2018), o Método RURUCAg visa melhorar a interface e as funcionalidades da aplicação para o usuário final coletando seus feedbacks e comentários, além de avaliar a qualidade da interface referente a comunicação do designer com os usuários de uma característica específica.

4.4 PROGRESSIVE WEB APPS (PWA)

Para superar os problemas existentes entre aplicações web, que exigem um navegador e uma boa conexão para abrir as aplicações e aplicativos nativos, que necessitam de um provedor de aplicativos como a Google PlayStore, foi criado o método de desenvolvimento PWA pela Google (KURNIAWAN, 2020). Para Sharma *et al.* (2019), utilizar o PWA implica em um baixo tempo de desenvolvimento e de implantação da aplicação, além de permitir que a empresa não contrate funcionários especializados em diferentes plataformas.

O PWA pode ser uma possível solução no problema descrito por Phuong (2019), no qual detalha que 90% das aplicações móveis baixadas em dispositivos móveis são excluídas após o uso. A taxa de usuários que visitam um site ultrapassa o pequeno número de aplicações móveis que os usuários instalam em seus dispositivos móveis, gerando assim uma perda financeira para a empresa se suas aplicações móveis são consideradas de uso único (PHUONG, 2019).

No contexto do desenvolvimento de aplicações móveis, há a necessidade de aprendizado em ambientes totalmente diferentes, resultando em um trabalho desafiador e pouco produtivo para atender diversas plataformas e versões de sistemas operacionais (SILVA; TIOSSO, 2020). O conjunto de padrões que o PWA proporciona, defendido pelo grupo Google Web Fundamentals, busca preencher as lacunas encontradas entre aplicações web e móveis introduzindo recursos como: suporte off-line, sincronização em segundo plano e instalação da aplicação web na tela inicial do dispositivo móvel (BIØRN-HANSEN *et al.*, 2017).

REFERÊNCIAS

ADAM, Noor L.; AIMAN SULAIMAN, Muhamad S.; SOH, Shaharuddin C. **Calculus video recommender system**. In: Journal of Physics: Conference Series 2019.

ALBERTUS, Michael; MULIADY, Muliady. **Pengaturan Fan Speed dan Suhu Air Conditioner Melalui Ucapan Dengan Layanan Google Assistant API**. TESLA: Jurnal Teknik Elektro, [S. l.], v. 21, n. 2, p. 170, 2020.

AZKYA, Zahra S.; ARDIANSAH, Irfan; PUJANTO, Totok. **Analisis User Experience pada Warehouse Marketplace dengan Metode Heuristic Evaluation**. Jurnal Teknik Informatika dan Sistem Informasi, [S. l.], v. 6, n. 1, 2020.

BAILON, Mark R. M. et al. **Filipino to chinese speech-to-speech translator using neural network with database system**. International Journal of Emerging Trends in Engineering Research, [S. l.], v. 7, n. 9, p. 276–282, 2019.

BARDAKCI, Salih. **Exploring High School Students' Educational use of Youtube**. International Review of Research in Open and Distance Learning, [S. l.], v. 20, n. 2, p. 260–278, 2019.

BARROZO, Ligia V.; SERAFIM, Mirela B.; MORAES, Sara L. de; MANSUR, Giselle. **Monitoramento Espaço-Temporal das Área de Alto Risco de COVID-19 nos Municípios do Brasil**. Hygeia - Revista Brasileira de Geografia Médica e da Saúde, [S. l.], p. 417–425, 2020.

BIØRN-HANSEN, Andreas; MAJCHRZAK, Tim A.; GRØNLI, Tor M. **Progressive web apps: The possibleweb-native unifier for mobile development**. In: WEBIST 2017 - Proceedings of the 13th International Conference on Web Information Systems and Technologies 2017, Anais [...]. [s.l.: s.n.].

BROOKS, Richard. **11 Google Translate Facts You Should Know**. K International, 02 mai. 2016. Disponível em: <https://k-international.com/blog/google-translate-facts/>. Acesso em: 14 out. 2020.

COSTA, Simone E. da. **iLibras como facilitador na comunicação do surdo: Desenvolvimento de um recurso colaborativo de tecnologia assistiva**. 2018. 263 f. Dissertação (Mestrado) – Centro de Ciências Tecnológicas, Universidade do Estado de Santa Catarina, 2018.

DESLANDES, Suely F.; COUTINHO, Tiago. **The intensive use of the internet by children and adolescents in the context of COVID-19 and the risks for self-inflicted violence.** *Ciência e Saúde Coletiva*, [S. l.], v. 25, p. 2479–2486, 2020.

DOHERTY, Stephen. **The impact of translation technologies on the process and product of translation.** *International Journal of Communication*, [S. l.], v. 10, 2016.

GOOGLE. Material Design. **Introduction.** [S.l.], 2019. Disponível em: <https://material.io/design/introduction/>. Acesso em: 22 nov. 2020.

HAMID, Saima; MIR, Mohammad Y.; ROHELA, Gulab K. **Novel coronavirus disease (COVID-19): a pandemic (epidemiology, pathogenesis and potential therapeutics).** *New Microbes and New Infections*, 2020.

IBGE. **Acesso à internet e à televisão e posse de telefone móvel celular para uso pessoal: 2016.** Rio de Janeiro: IBGE, 2018.

KHECHINE, Hager; LAKHAL, Sawsen. **Technology as a double-edged sword: From behavior prediction with UTAUT to students' outcomes considering personal characteristics.** *Journal of Information Technology Education: Research*, [S. l.], v. 17, p. 63–102, 2018.

KURNIAWAN, Antonius A. **Analisis Performa Progressive Web Application (PWA) Pada Perangkat Mobile.** *Jurnal Ilmiah Informatika Komputer*, [S. l.], v. 25, n. 1, 2020.

LAKOMKIN, Egor; MAGG, Sven; WEBER, Cornelius; WERMTER, Stefan. **KT-Speech-Crawler: Automatic dataset construction for speech recognition from YouTube videos.** *EMNLP 2018 - Conference on Empirical Methods in Natural Language Processing: System Demonstrations, Proceedings 2018*.

MAIA, Maria A. Q.; BARBOSA, Ricardo R.; WILLIAMS, Peter. **Usabilidade e experiência do usuário de sistemas de informação: em busca de limites e relações.** *Ciência da Informação em Revista*, [S. l.], v. 6, n. 3, 2020.

MALIK, Haroon; TIAN, Zifeng. **A Framework for Collecting YouTube Meta-Data.** In: *Procedia Computer Science 2017, Anais [...]*. : Elsevier B.V., 2017. p. 194–201.

NAGUMO, Estevon; TELES, Lúcio F.; SILVA, Lucélia de A. **A utilização de vídeos do Youtube como suporte ao processo de aprendizagem (Using Youtube videos to support the learning process).** *Revista Eletrônica de Educação*, [S. l.], v. 14, p. 3757008, 2020.

PHUONG, Nguyen H. **Progressive Web App in Enhancing App Experience: life care insurance claim application.** 2019. 82 f. TCC (Graduação) - Information Technology, Vaasan Ammattikorkeakoulu University Of Applied Sciences, Vaasa, 2019.

SHARMA, Varsha et al. **Progressive Web App (PWA) - One Stop Solution for All Application Development Across All Platforms.** *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, [S. l.], p. 1120–1122, 2019.

SILVA, Carlos E. da; FERNANDES, Lincoln P. **Introducing COPA-Trad version 2.0 a parallel corpus-based system for translation research, teaching and practice**. Ilha do Desterro, [S. l.], v. 73, n. 1, 2020.

SILVA, Jonathas K.; TIOSSO, Fernando. **Revisão Bibliográfica sobre Conceito de Progressive Web Applications (PWA)**. Revista Interface Tecnológica, [S. l.], v. 17, n. 1, p. 53–64, 2020.

TORRES, Cecília V. et al. **DESENVOLVIMENTO MOBILE COM ENFOQUE ACESSÍVEL: O DESIGN NA MEDIAÇÃO DA INCLUSÃO**. Human Factors in Design, [S. l.], v. 7, n. 13, p. 85–101, 2018.

Comentado [AS1]: Escrever em caixa baixa.

UKO, Victor S.; DUBUKUMAH, Gachada B.; AYOSUBOMI, Ibrahim K. **Microcontroller Based Speech to Text Translation System**. European Journal of Engineering Research and Science, [S. l.], v. 4, n. 12, p. 149–154, 2019.

XAVIER, Marcelle B.; CARVALHO, Francisco dos S.; CARVALHO, Mauro dos S.; MORAES, Juliana M. de. **Identidade Surda: Uma Análise dos Discursos de Professores Surdos e Ouvintes no Youtube / Deaf Identity: An Analysis of the Discourses of Deaf Teachers and Listeners on Youtube**. ID on line REVISTA DE PSICOLOGIA, [S. l.], v. 13, n. 45, 2019.

ASSINATURAS

(Atenção: todas as folhas devem estar rubricadas)

Assinatura do(a) Aluno(a): _____

Assinatura do(a) Orientador(a): _____

Assinatura do(a) Coorientador(a) (se houver): _____

Observações do orientador em relação a itens não atendidos do pré-projeto (se houver):

FORMULÁRIO DE AVALIAÇÃO – PROFESSOR TCC I

Acadêmico(a): Hugo Marcel Larsen _____

Avaliador(a): Andreza Sartori _____

ASPECTOS AVALIADOS ¹		atende	atende parcialmente	não atende
ASPECTOS TÉCNICOS	1. INTRODUÇÃO O tema de pesquisa está devidamente contextualizado/delimitado?	X		
	O problema está claramente formulado?	X		
	2. OBJETIVOS O objetivo principal está claramente definido e é passível de ser alcançado?	X		
	Os objetivos específicos são coerentes com o objetivo principal?	X		
	3. JUSTIFICATIVA São apresentados argumentos científicos, técnicos ou metodológicos que justificam a proposta?	X		
	São apresentadas as contribuições teóricas, práticas ou sociais que justificam a proposta?	X		
	4. METODOLOGIA Foram relacionadas todas as etapas necessárias para o desenvolvimento do TCC?	X		
ASPECTOS METODOLÓGICOS	Os métodos, recursos e o cronograma estão devidamente apresentados?	X		
	5. REVISÃO BIBLIOGRÁFICA (atenção para a diferença de conteúdo entre projeto e pré-projeto) Os assuntos apresentados são suficientes e têm relação com o tema do TCC?	X		
	6. LINGUAGEM USADA (redação) O texto completo é coerente e redigido corretamente em língua portuguesa, usando linguagem formal/científica?	X		
	A exposição do assunto é ordenada (as ideias estão bem encadeadas e a linguagem utilizada é clara)?	X		
	7. ORGANIZAÇÃO E APRESENTAÇÃO GRÁFICA DO TEXTO A organização e apresentação dos capítulos, seções, subseções e parágrafos estão de acordo com o modelo estabelecido?	X		
	8. ILUSTRAÇÕES (figuras, quadros, tabelas) As ilustrações são legíveis e obedecem às normas da ABNT?	X		
	9. REFERÊNCIAS E CITAÇÕES As referências obedecem às normas da ABNT?	X		
	As citações obedecem às normas da ABNT?	X		
	Todos os documentos citados foram referenciados e vice-versa, isto é, as citações e referências são consistentes?	X		

PARECER – PROFESSOR DE TCC I OU COORDENADOR DE TCC (PREENCHER APENAS NO PROJETO):

O projeto de TCC será reprovado se:

- qualquer um dos itens tiver resposta NÃO ATENDE;
- pelo menos **4 (quatro)** itens dos **ASPECTOS TÉCNICOS** tiverem resposta ATENDE PARCIALMENTE; ou
- pelo menos **4 (quatro)** itens dos **ASPECTOS METODOLÓGICOS** tiverem resposta ATENDE PARCIALMENTE.

PARECER: (x) APROVADO () REPROVADO

Assinatura: _____ Data: 07/12/2020 _____

¹ Quando o avaliador marcar algum item como atende parcialmente ou não atende, deve obrigatoriamente indicar os motivos no texto, para que o aluno saiba o porquê da avaliação.