

PROJETO II - Última Entrega

Este artigo tem o objetivo documentar o projeto final curricular de Análise e Desenvolvimento de Sistemas da Faculdade Senac Porto Alegre.

Resumo do Projeto

DRIVER'S TEM" APP/API - codificar voz em texto para maior segurança no trânsito.

A digitalização modernizou e acelerou a conexão entre clientes e fornecedores e entre grupos de pessoas associadas para uma determinada atividade por equipamentos móveis, computadores, notebooks e aparelhos de celular. Entretanto, a digitalização também trouxe malefícios aos seus usuários, por exemplo, a atividade de motoristas de táxi e de transportes de mercadorias ganhou um novo fator de risco com a popularização dos smartphones, por reduzir a atenção do motorista ao volante quando uma nova mensagem é apresentada nestes equipamentos móveis.

As paradas dos veículos (carro ou moto) para responder com segurança aos seus clientes por mensagens de texto causa atrasos no decorrer dos serviços prestados. A eventual irresponsabilidade em responder mensagens durante o percurso pode causar acidentes.

O aplicativo "DRIVER'S TEM" visa transformar a voz de formato áudio para formato de mensagens de texto integrando outros aplicativos. Com o "DRIVER'S TEM" os motoristas não mais terão a necessidade de digitar mensagens, manterão sua atenção ao volante enquanto suas mensagens de voz são interpretadas e enviadas ao destinatário rapidamente.

Definição do Problema

Companhias de Táxi ainda utilizam sistemas de radiofrequência como alternativa mais segura para taxistas associados e sob inscrição na área de telecomunicações de âmbito federal (órgão responsável ANATEL). Este sistema inclui o registro de cada estação, ou seja, para cada veículo tipo táxi há uma taxa anual de serviço de telecomunicação nacional. A troca de informações via voz entre o motorista do veículo e a central de atendimento captura as informações e as repassa nos sistemas pertinentes em momentos diferentes.

(<https://www.gov.br/pt-br/servicos/obter-autorizacao-para-servico-limitado-privado>)

Manter o sistema de rádio frequência requer um investimento anual entre R\$ 400,00 e R\$ 1200,00 para o motorista de táxi. Com a aplicação de telefone celular, Google Maps e outros aplicativos que podem "rodar" em uma só plataforma, vários motoristas estão mudando adesão a novas companhias que promovem uso de aplicativos via web/cloud com troca de mensagens de texto sem custo adicional. No entanto, estes aplicativos podem gerar distração ao volante e aumentar riscos de acidentes de trânsito.

Segundo os registros estatísticos da NHTSA (Administração Nacional de Segurança Rodoviária dos Estados Unidos), distrações com mensagens de texto são seis vezes mais perigosas que dirigir após ingerir bebida alcoólica.

(<https://fenasdetran.com/noticia/distracao-ao-volante>)

Uma pesquisa do VTTI (Instituto de Transportes e Tecnologia da Virgínia), órgão ligado à NHTSA, demonstrou as seguintes periculosidades geradas por mensagens de texto no decorrer do percurso com distração manual, visual e cognitiva ao mesmo tempo:

- LER MENSAGENS DE TEXTO
 - Tempo de desatenção = 2 segundos
 - Distância percorrida a 60 km/h: 34 m, a 100 km/h: 56 m
- ESCREVER MENSAGEM DE TEXTO
 - Tempo de desatenção = 2,5 segundos
 - Distância percorrida a 60 km/h: 42 m, a 100 km/h: 69 m

A segurança do veículo, motorista e passageiros estará em alto risco caso o motorista irresponsavelmente tente ler e responder mensagens de texto ao aplicativo de táxi, ou de entrega de mercadorias para dar informações de localização, ou horários ou disponibilidade. Ainda, as “corridas” poderão sofrer atrasos caso o motorista, desta vez responsavelmente, decida parar o veículo de forma segura para ler e responder mensagens. Ou seja, em ambas situações, de risco ou de paradas, há falta de praticidade para a atividade de dirigir. A usabilidade do sistema antigo de chamadas de rádio com trocas de informações por voz ainda é a mais segura e prática.

Este projeto tem o objetivo de unir a modernização da digitalização através da segurança e praticidade da aplicação do formato da informação em voz (áudio) que poderá ser convertido em formato texto por um aplicativo (APP/API) de conversão e interpretação para integrar a informação com outros aplicativos.

Análise de Sistemas Correlatos

Suporte nativo a digitação por voz (Google)

Este recurso já vem instalado por padrão nos celulares Android. Permite inserir diretamente o texto ditado no aplicativo onde a mensagem será enviada. Não possui limite de uso. Não permite personalização, insere exatamente como o texto é falado e não insere pontuação.

Escrever por voz (UXApps Ltd.)

Aplicativo simples de usar, só abrir e sair usando, possuindo opção de compartilhar a mensagem transcrita ao lado do botão de gravar. Bastante intuitivo. Versão grátis com uso ilimitado. Apresenta alternativas da frase gravada, com correções da linguagem coloquial, mas a pontuação nem sempre estava correta nos testes feitos e não possui possibilidade de customização.

Transcriptor

Aplicativo complexo de usar, após gravar um áudio para ser transcrito é preciso voltar à tela anterior e entrar em outra área para acessar a transcrição. O processo de transcrição se mostrou sempre bastante demorado em comparação com os outros. A versão gratuita tem limitação de uso de 5 minutos de áudio gravado por dia.

Objetivos

O projeto tem como objetivo geral, desenvolver um APP Web/Mobile (como piloto inicialmente em formato de API) convertendo voz em texto, que atenda as necessidades do usuário (motorista de veículo exercendo trabalho de transporte de passageiros ou de mercadorias) sobre a troca de informações de horários, localizações geográficas, situações de trânsito, etc. para informar os seus clientes.

Objetivos Específicos

Este aplicativo tipo API interface (como piloto) possibilitará a interação com a máxima segurança no trânsito como principal adequação da digitalização ao ambiente de trabalho de motoristas de veículos por utilização de voz, e convertendo este formato áudio em texto para maior facilidade e rapidez de integrá-las às plataformas de computação WEB das empresas de associações de táxi e empresas de entrega de mercadorias e, consequentemente, em comunicação com clientes que estão buscando o serviço de transporte. Inicialmente construímos um exemplo “piloto” na plataforma Android e que poderá ser adaptada para outros tipos de sistemas, bem como uma aplicação servidor para fazer a conversão do áudio em voz. Os Objetivos específicos, portanto, são:

- Desenvolver um aplicativo Android para gravar áudios de motoristas, enviar ao servidor e compartilhar o texto recebido como resposta com outros aplicativos do aparelho móvel;
- Desenvolver uma aplicação servidor para receber os áudios do aplicativo Android e fazer a conversão para texto;
- Configurar um banco de dados para armazenar textos que serão usados como tipos possíveis ao fazer uma conversão de áudio em texto.

Stack Tecnológico

Front-end

Cordova

Como está descrito no site do projeto, Apache Cordova é um framework para desenvolvimento de aplicações mobile de código aberto. A sua vantagem é que o aplicativo é programado uma única vez utilizando HTML e JavaScript, mas a mesma solução pode ser compilada tanto para Android quanto para IOS. Em teoria seria possível usar o mesmo código para criar uma aplicação WEB também, mas durante a utilização nesse trabalho notou-se que a maioria dos plugins não tem suporte para essa terceira opção.

JavaScript

Segundo David Flanagan, em JavaScript - O Guia Definitivo, essa é a linguagem de programação mais onipresente da história. É uma linguagem interpretada estruturada, de script em alto nível com tipagem dinâmica fraca e multiparadigma. Juntamente com HTML e CSS, o JavaScript é uma das três principais tecnologias da World Wide Web.

Back-end

Oracle Database

Oracle Database é um sistema gerenciador de banco de dados. Como é explicado em Oracle Essentials, de Rick Greenwald et al, foi o primeiro banco de dados relacional a ser disponibilizado no mercado, sendo até hoje referência nessa área. Nesse trabalho, utilizaremos a Oracle Cloud, uma plataforma que permite utilização otimizada e aberta para conexão de qualquer lugar do mundo.

Python

Segundo Luiz Borges, em Python para desenvolvedores, Python é uma linguagem multiparadigma mantida pela Python Software Foundation. Possui suporte a orientação a objetos e mesmo os tipos básicos são objetos. É usado como linguagem principal no desenvolvimento de sistemas, mas como suporta a execução de comandos unitários sem a necessidade de compilação, é muito usado como linguagem de script.

Google API Speech-to-text

Google Cloud é uma plataforma de diversas bibliotecas para fins específicos de transformação de dados através de APIs/interfaces. Entre estas bibliotecas a Google desenvolveu o API Speech-to-text, que permite receber um formato arquivo áudio e transformá-lo em formato texto e importante notar que esta transformação inclui reconhecimento de múltiplas línguas, desde ingles, a portugues e francês e demais línguas atualmente já suportadas em outros serviços e aplicativos da Google como o Tradutor Google. A aplicação deste API em nosso projeto cria antecipadamente a possibilidade de aplicação em vários países mundialmente.

<https://cloud.google.com/text-to-speech>

Descrição da Solução

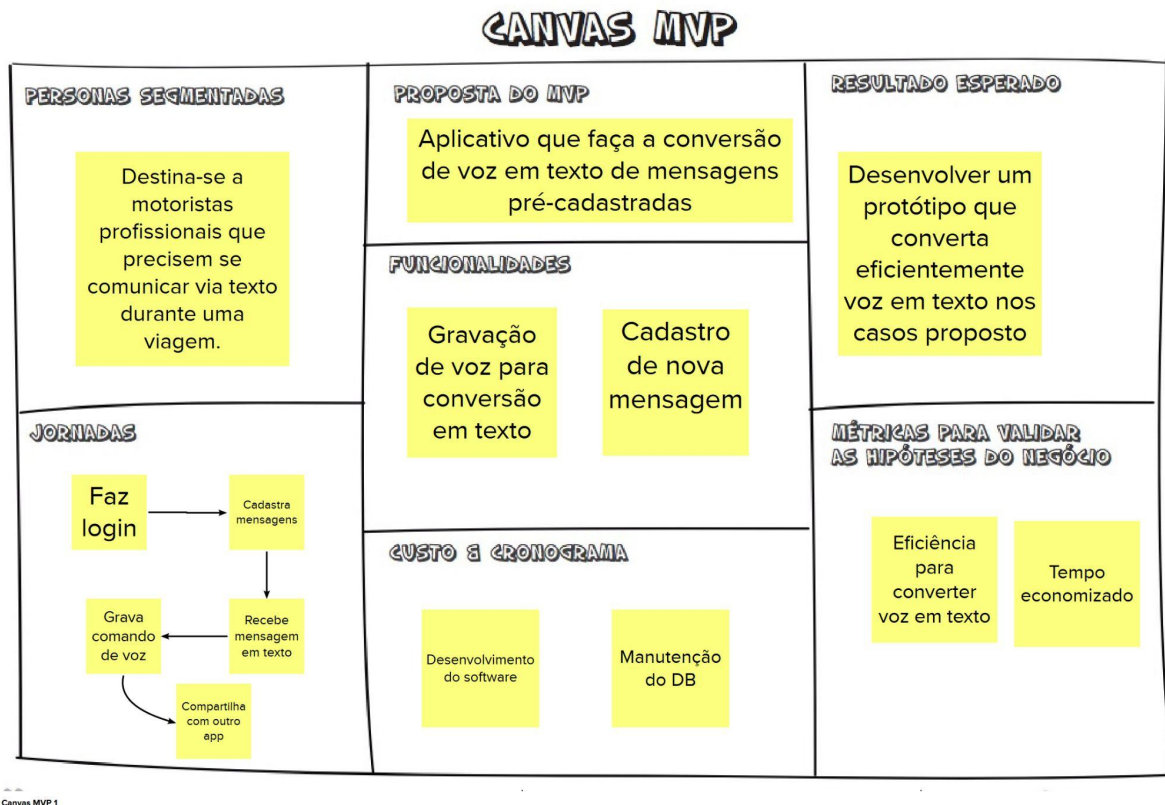
Para atender aos objetivos citados, implementamos uma solução composta por um aplicativo Android e uma API servidor que se comunicam via HTTP. O aplicativo para Android é responsável por capturar as mensagens de áudio gravadas pelo motorista e enviar para o servidor. A aplicação servidor faz a conversão do áudio recebido em texto e enviar de volta para o aplicativo Android.

O aplicativo para Android, que foi desenvolvido em HTML e Javascript, conta com uma opção para gravar uma mensagem de áudio para ser convertida em texto. Ao receber a resposta do servidor, o texto recebido pode ser compartilhado com outros aplicativos. Também foi implementada uma opção para cadastrar novos tipos de mensagens de texto possíveis.

A aplicação servidor, que foi implementada em Python, ao receber uma mensagem de áudio enviada pelo aplicativo Android, usará a Google Speech API para fazer a conversão para texto. É previsto que o texto gerado pela Google Speech API não tenha sido convertido perfeitamente, por isso o resultado dessa conversão é comparado com os tipos previamente armazenados em um banco Oracle e é escolhido o texto mais parecido com o convertido, tendo um percentual mínimo de semelhança para ser considerado válido. Ao receber uma mensagem de texto enviada pelo aplicativo Android, adicionará aos tipos de texto possíveis armazenados no banco Oracle.

Arquitetura

Canvas



Personas

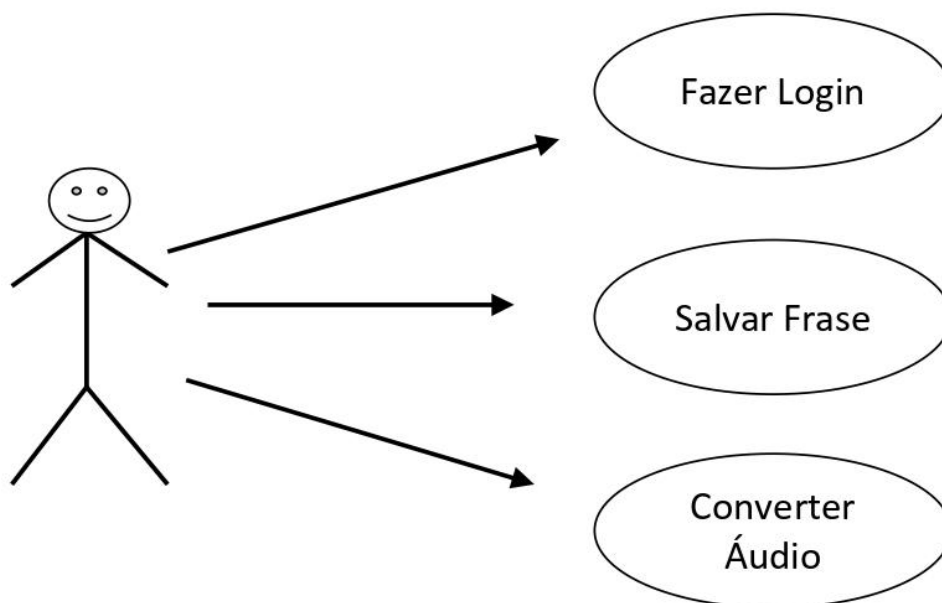
1. Celso (62 anos)
Taxista há 40 anos. Durante muito tempo usou apenas rádio para se comunicar com o ponto que recebe chamados dos clientes. Nos últimos 5 anos começou a usar EasyTaxi para pegar corridas também, mas tem dificuldade em digitar no celular. Gostaria de uma solução que o poupasse de digitar tanto.
2. Luana (35 anos)
Motorista de uber. Formada em administração, começou a trabalhar como motorista há 2 anos para complementar a renda e acabou virando sua ocupação principal. Usa o aplicativo da Uber para se comunicar com os clientes e o WhatsApp para se comunicar com familiares. Se preocupa em responder rapidamente tanto seus clientes para ser bem avaliada quanto seus familiares para tranquilizá-los, já que se preocupam com sua segurança.
3. João (41 anos)

Motorista de caminhão contratado há 10 anos. Usa SMS para se comunicar com a empresa. Manda sempre as mesmas mensagens padrão e gostaria de automatizar essa tarefa.

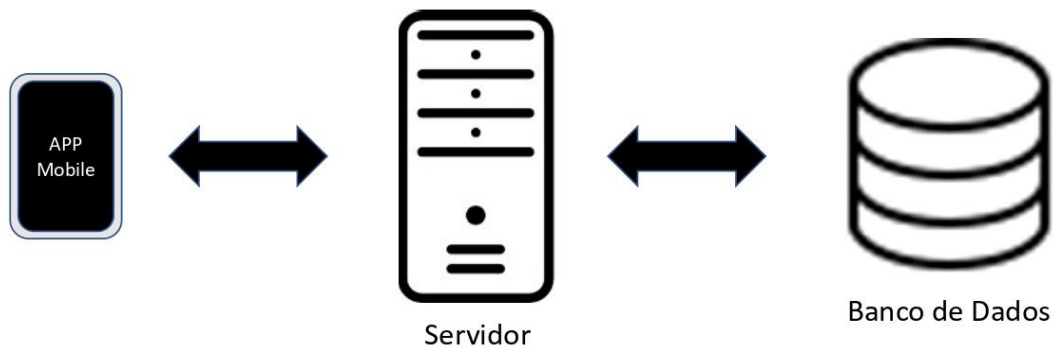
Benchmarking

	Suporte nativo	Escrever por voz	Transcriptor	DRIVER'S TEM
Usabilidade	Muito fácil de usar	Fácil de usar	Difícil de usar	Fácil de usar
Acesso	Nativo	Livre	Parcial	Livre
Customização	Não possui	Média	Não possui	Alta
Acuidade	Boa	Boa	Boa	Perfeita

Casos de uso



Representação da arquitetura



Repositório do Projeto

Todos os arquivos da solução, tanto a parte mobile quanto servidor foram disponibilizadas no repositório GIT: https://github.com/Riccardoneves1976/projetos_II.git

Validação

Estratégia

Para verificar se a solução proposta atende aos objetivos listados serão feitos dois tipos de validação: sobre a funcionalidade do sistema e sobre o atendimento ao objetivo geral de reduzir o tempo de resposta de mensagens.

- Funcionalidade do sistema
 - Para verificar se o sistema está cumprindo seu objetivo de converter áudio em texto serão gravadas no aplicativo Android algumas frases que potencialmente seriam utilizadas pelos usuários (listadas abaixo). Seria anotado se o sistema consegue fazer a conversão corretamente.
- Atendimento ao objetivo geral de reduzir o tempo de resposta de mensagens
 - Para verificar se o sistema atende ao objetivo geral de reduzir o tempo de resposta por escrito, serão feito testes analisando o tempo ativo (ou seja, o tempo em que a atenção do usuário é requerida no aplicativo, não incluindo o tempo de espera pela resposta do servidor, por exemplo) requerido para uma resposta utilizando o aplicativo e utilizando a inserção de texto convencional. Para este segundo teste serão usadas as mesmas frases listadas.

Consolidação de dados

- Funcionalidade do sistema
 - Os dados colhidos fazendo o teste descrito serão apresentados na forma de um gráfico de pizza mostrando o percentual de acerto para cada tipo de frase testado e o percentual geral encontrado.
- Atendimento ao objetivo geral de reduzir o tempo de resposta de mensagens

- Os dados colhidos neste teste serão apresentados gráficos de barras comparativos entre o tempo gasto usando o aplicativo e usando a digitação manual para cada tipo de frase testado e o tempo médio geral.

Conclusões

Os objetivos específicos foram atingidos com sucesso, tendo sido desenvolvida uma aplicação mobile que captura áudio e envia para um servidor convertê-lo em texto. Após a conversão o seu resultado é enviado de volta para o cliente, que permite seu compartilhamento com outros aplicativos. Também é possível escolher quais frases serão candidatas à conversão.

Limitações do Projeto e Perspectivas Futuras

Em razão dos tempo gasto muito superior ao estimado na fase de configuração do ambiente de desenvolvimento, ou seja, instalação de todas as ferramentas e dependências necessárias, não foi possível fazer os testes quantitativos com o rigor que se pretendia. Qualitativamente notou-se que a qualidade do áudio captado estava muito pior do que acreditava-se que seria alcançado, o que fez com que a usabilidade do aplicativo fosse afetada.

Ainda que o áudio gravado de maneira como um usuário comum faria seja compreensível, para que a conversão seja satisfatória é preciso que o se fale alto e articuladamente e não deixe espaços silenciosos no início e fim da gravação. Entretanto, seguindo esses passos, praticamente todas as conversões testadas tiveram sucesso. Para uma próxima versão seria bom testar outros métodos de aquisição de áudio, além do plugin Cordova que foi utilizado.

Ainda devido à falta de tempo, não foi implementada a tela de login, que permitiria a utilização do aplicativo por múltiplos usuários. Em uma versão futura as frases salvas no banco estão atreladas a um ID criado para cada usuário.

Também seria desejado melhorar a aparência geral do aplicativo, o que acabaria beneficiando também sua usabilidade, já que nessa primeira fase optou-se por priorizar a funcionalidade proposta.

Referências Bibliográficas

https://books.google.com.br/books?hl=pt-BR&lr=&id=oSKwAAAAQBAJ&oi=fnd&pg=PT3&dq=oracle+database&ots=wRqsW1RhLJ&sig=Ri4ZiDAzVbvvnH7zcLArt-DAgJ8&redir_esc=y#v=onepage&q=oracle%20database&f=false

David Flanagan, em JavaScript - O Guia Definitivo
(https://books.google.com.br/books?hl=pt-BR&lr=&id=zWNYDgAAQBAJ&oi=fnd&pg=PR1&dq=javascript&ots=IAAiW4KajI&sig=w42X9-8OTyXJesbXZEPRLOE5Mwg&redir_esc=y#v=onepage&q=javascript&f=false)

https://books.google.com.br/books?hl=pt-BR&lr=&id=eZmtBAAQBAJ&oi=fnd&pg=PA14&dq=python&ots=VERunoEkeo&sig=Ec2qnoYE9zUIZOZeQdbAqFa2wlg&redir_esc=y#v=onepage&q=python&f=false

<https://cloud.google.com/speech-to-text>

<https://caroli.org/cartazes-lean-inception/>

<https://app.mural.co/t/mycomp8113/m/mycomp8113/1666964409532/c6099426802577bd64d0b3b890cd011527196214?sender=u4f46f5020437227739486095>

<https://cordova.apache.org/docs/en/latest/guide/overview/index.html>

<https://www.npmjs.com/package/cordova-plugin-advanced-http>

<https://www.npmjs.com/package/cordova-plugin-media>

<https://www.npmjs.com/package/cordova-plugin-x-socialsharing>

<https://medium.com/@everton.tomalok/calculando-similaridades-entre-strings-ebbea21d5b7a>

[1] Borges, Luiz Eduardo. *Python para desenvolvedores: aborda Python 3.3*. Novatec Editora, 2014.