

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO**

**Um estudo experimental investigando  
soluções text-to-speech**

**Pilar Baptista Fernandez**

**PROPOSTA DE PROJETO FINAL DE GRADUAÇÃO**

**CENTRO TÉCNICO CIENTÍFICO – CTC**

**DEPARTAMENTO DE INFORMÁTICA**

**Curso de Graduação em Ciência da Computação**

Rio de Janeiro, Outubro de 2020

## **1. Introdução**

Existem diversas soluções de text-to-speech (Sasirekha, D., and E. Chandra, 2012) disponíveis. Estas tecnologias buscam reproduzir artificialmente a fala humana, se baseando no processamento de textos escritos fornecidos. A síntese de fala tem uma função importante de aumentar a acessibilidade à conteúdos textuais, permitindo que pessoas analfabetas ou com deficiência visual compreendam facilmente o que está escrito, mesmo não conseguindo ler.

Entretanto, é difícil obter uma visão geral das estratégias utilizadas em cada uma dessas soluções e da sua eficácia para a comunicação com pessoas. Atualmente, não há um apoio para auxiliar desenvolvedores de software na tomada de decisão em relação a qual solução de text-to-speech se adequa melhor ao que deseja implementar e ao seu público alvo ou usuário final. Este apoio deveria compreender conhecimento sobre o esforço para implementação e também a eficácia das diferentes soluções.

## **2. Proposta e Objetivos do trabalho**

No contexto deste trabalho será provida a fundamentação teórica de como funcionam soluções text-to-speech.

O objetivo deste trabalho é:

- (i) Fazer um levantamento de soluções text-to-speech e uma descrição das estratégias utilizadas em cada uma;
- (ii) Conduzir um estudo experimental comparando as soluções em relação ao esforço de desenvolvimento e à eficácia da comunicação.

Com os resultados e o conhecimento gerado neste estudo pretendemos auxiliar desenvolvedores de software que estejam procurando utilizar alguma destas tecnologias no desenvolvimento de suas aplicações.

## **3. Plano de Ação**

Para o levantamento será realizada uma revisão de artigos da literatura. Com base nesta revisão e em buscas adicionais, serão identificadas soluções relevantes de text-to-speech. Estas soluções serão investigadas com pequenos

exemplos de esforço e eficácia, orientações para desenvolvedores serão providas. implementação refletindo sobre o esforço para sua utilização e seu desempenho. Na sequência, um estudo experimental (Wohlin et al., 2012) sobre a eficácia de comunicação destas soluções será conduzido. Com base em uma discussão sobre as informações sobre esforço e eficácia, orientações para desenvolvedores serão providas.

Dito isso, o cronograma proposto para a realização do projeto completo é:

- 1- Realizar revisão da literatura sobre text-to-speech e escrever sobre a fundamentação teórica: De outubro/2020 a dezembro/2020;
- 2- Identificar soluções relevantes: De outubro/2020 a dezembro/2020;
- 3- Implementar exemplos com as soluções encontradas, refletindo sobre esforço: março/2021;
- 4- Realizar um estudo experimental sobre a eficácia de comunicação das soluções: março/2021 a abril/2021;
- 5- Elaborar orientações para desenvolvedores: De abril/2021 a junho/2021;
- 6- Finalizar escrita da monografia e defender projeto final: junho ou julho/2021.

## 4. Referências bibliográficas

Sasirekha, D., Chandra, E. **Text to speech: a simple tutorial**. International Journal of Soft Computing and Engineering (IJSCE).  
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.682.5362&rep=rep1&type=pdf>, 2012.

WOHLIN, Claes *et al.* **Experimentation in software engineering**. Springer Science & Business Media, 2012.