

AVALIAÇÃO 1 - ANALISADOR LÉXICO

A etapa de análise léxica inclui as seguintes atividades para confecção do seu tradutor/compilador/chatbot:

- 1) Verificação léxica, com **avaliação do alfabeto usado**. Basicamente **aceite letras de A-Z, a-z, 0-9, caracteres de pontuação e caracteres especiais presentes no teclado**. Ignore caracteres especiais, visíveis ou invisíveis.
- 2) Como se trata de linguagem natural, providencie **uma lista de stopwords** que possa ser mobilizada em seu chatbot. A lista pode ser obtida na Internet, de qualquer fonte (indique a fonte). Observe que as palavras-chaves usadas em seu trabalho anterior podem exigir a mudança da lista de stopwords. Altere-a.
- 3) Geração de lexemas é a redução de palavras encontradas no texto do seu usuário. Nosso chatbot usa Português Brasileiro. **Você pode usar um método pronto de geração de lexemas**. Obtenha um (indique a fonte) ou crie um.
- 4) Todas as palavras presentes no texto do seu usuário que não sejam palavras-chaves e stopwords devem ser incluídos em uma **estrutura de dados do tipo lista**, sua Tabela de Símbolos. Implemente sua Tabela de Símbolos preliminar.

Valor: 25 pontos

Data de entrega: 24/05/2022.

Aluno: Tássis Fernando Soares Fernandes

Referências:

- Lista de Stopwords:
https://github.com/thiagoscouto/stopwords_ptbr/blob/master/stopwords_ptbr.txt
- Biblioteca de geração de lexemas:
<http://snowball.tartarus.org/download.html>