

## **Inteligência Artificial**

### **PLN - Biblioteca spaCy**

### **Projeto – 2º bimestre**

Fazer um exercício em Python, com uso da biblioteca spaCy, para as extrações abaixo. A entrada para as análises de PLN será através de um mesmo arquivo texto. A saída gerada também será um único arquivo texto. Para a similaridade entre tokens, o programa deverá solicitar duas palavras de entrada e efetuar a análise. Acrescentar na saída a quantidade de ocorrência das duas palavras no arquivo de entrada. Na saída, identificar as palavras utilizadas.

- Tokenização
- Part-of-Speech Tagging (POS)
- Lematização
- Stop Words
- Reconhecimento de Entidades Nomeadas (NER)
- Segmentação de Sentenças
- Morfologia
- Similaridade entre Tokens