



UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO  
CENTRO DE EDUCAÇÃO E HUMANIDADES  
INSTITUTO DE LETRAS  
COORDENAÇÃO DE PÓS-GRADUAÇÃO EM LETRAS

**PROGRAMA DE DISCIPLINA 2019.2**

<b>Área</b>	( x ) Estudos de Língua ( ) Estudos de Literatura
<b>Especialidade</b>	( ) Língua Portuguesa ( ) Literatura Brasileira ( x ) Linguística ( ) Literatura Portuguesa ( ) Literaturas de Língua Inglesa ( ) Teoria da Literatura e Literatura Comparada
<b>Nível</b>	( x ) Mestrado ( x ) Doutorado

<b>Disciplina</b>	<b>Estudos em linguística e ciências afins</b>
<b>Tema</b>	Sociolinguística e Processamento de Linguagem Natural: explorações e práticas
<b>Professor(a)</b>	Ricardo Joseh Lima
<b>Dia e horário</b>	<b>5ª feira, 8h50 às 12h20</b>
<b>Recursos audiovisuais</b>	( X ) Sim ( ) Não ( ) Eventualmente

**Ementa**

A Sociolinguística Variacionista. Processamento de Linguagem Natural (PLN). Ferramentas de PLN. Aplicações práticas.

**Programa**

O curso será oferecido em formato de oficinas. A parte prática acontecerá inicialmente no espaço da aula. Os textos em inglês não constituem leitura obrigatória e são apenas referências para consulta.

1. Sociolinguística
  - 1.1 O estudo da variação
  - 1.2 O aspecto quantitativo
  - 1.3 Ferramentas estatísticas
2. Processamento de Linguagem Natural
  - 2.1 A abordagem computacional para a linguagem
  - 2.2 O aspecto quantitativo
  - 2.3 Ferramentas computacionais: Python e suas bibliotecas (expressões regulares, etiquetadores gramaticais)
3. Aplicações práticas
  - 3.1 Busca de palavras em grandes quantidades de textos
  - 3.2 Classificação gramatical para fins didáticos
  - 3.3 Corretor gramatical: limites e possibilidades

### Bibliografia Inicial

Avanço, L. **Sobre normalização e classificação de polaridade de textos opinativos na web.** Dissertação de mestrado : Universidade de São Paulo, 2015.

Connelly, E., del Genio, C., Harrison, F. Datamining a medieval medical text reveals patterns in ingredient choice that reflect biological activity against the causative agents of specified infections. Disponível em <https://arxiv.org/abs/1807.07127v1>, 2018.

Ferreira, M., Lopes, M. **Para conhecer linguística computacional.** Contexto : São Paulo, 2019.

Higuchi, S., Freitas, C., Cuconato, B., Rademaker, A. Text Mining for History: first steps on building a large dataset. Disponível em <http://cpdoc.fgv.br/acervo/dhbb>, 2018.

Jargas, A. **Expressões regulares: uma abordagem divertida.** Novatec : São Paulo, 4<sup>a</sup> ed., 2012.

Nguyen, D., Doğruöz, A., Rosé, C., de Jong, F. Computational sociolinguistics: A survey. Computational Linguistics, 42(3):537–593, 2016.