



# Normalização de Texto

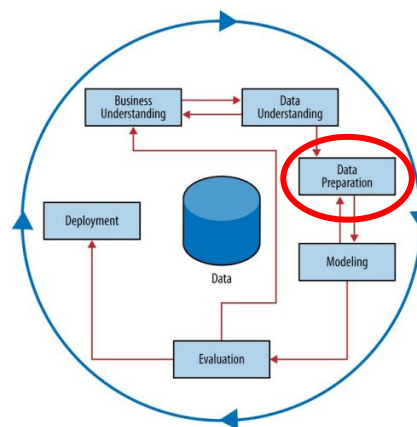
*Prof. Marcelo Pita*

Pós em Inteligência Artificial  
Processamento de Linguagem Natural



# Normalização de texto

**Normalização** é o conjunto de rotinas de tratamento e transformação aplicadas sobre o texto a fim de deixá-lo adequado para a modelagem analítica.



# Normalização de texto

---

Rotinas que vamos explorar:

- Transformação para minúsculas
- Remoção de palavras por tamanho
- Remoção de *stop words*.
- Remoção de sinais de pontuação
- Estemização e lematização
- Remoção de sinais de acentuação
- Remoção de palavras por frequência