

Organização e Recuperação da Informação

Wendel Melo

Faculdade de Computação
Universidade Federal de Uberlândia

Organização e Recuperação da Informação

Apresentação

Apresentação

- **Formação:**
 - Cientista Computacional (UFRJ - *Bsc, Msc, Dsc*)
- **Interesses:**
 - Otimização computacional
 - Inteligência computacional
 - Computação científica
- **Linguagens de Programação prediletas:**
 - C++
 - Python

Apresentação

- **Contato:**

- wendelmelo@ufu.br
- <https://wendelmelo.net>

- **Atendimento:**

- Terças e Quintas, online:
 - Terças - 17:00 às 18:00
 - Quintas - 17:00 às 18:00
- Horários extras (presenciais ou online) podem ser agendados

Bibliografia

- Ricardo Baeza-Yates, Berthier Ribeiro-Neto; Recuperação de Informação - Conceitos e Tecnologia das Máquinas de busca. Ed Bookman.
- Manning, Raghavan & Schutze; Introduction to information Retrieval.
- Wendel Melo; IUP: Introdução ao Universo da Programação com Python (em construção). <https://wendelmelo.net/book>

Avaliação

- Provas (70 Pontos):
 - Prova 1 (35 Pontos): 03/10/2024
 - Prova 2 (35 Pontos): ~~14/11/2024~~ 07/11/2024
 - Prova de recuperação (70 Pontos): ~~19/11/2024~~ 14/11/2024

Avaliação

- Provas (70 Pontos):
 - Prova 1 (35 Pontos): 03/10/2024
 - Prova 2 (35 Pontos): 14/11/2024
 - Prova de recuperação (70 Pontos): 19/11/2024
- **A prova de recuperação cobrirá a matéria toda! Sua nota substituirá a soma de notas da Prova 1 e da Prova 2;**
- É permitido faltar à Prova 1 ou à Prova 2 sem justificativa. Nesse caso, o(a) estudante deverá fazer a prova de recuperação para substituir a prova faltante (apenas cobrindo a matéria da prova faltante).

Avaliação

- 2 trabalhos **individuais** de implementação em Python (30 Pontos);
- 5 pontos extras por participação nas aulas.

Tópicos Principais

- Introdução;
- Recuperação de dados X Recuperação de Informação;
- Construção de sistemas de busca;
- Operações sobre a base de documentos e indexação;
- Modelos clássicos (booleano, vetorial e probabilístico);
- Avaliação da recuperação;
- Expansão de consultas;
- Agrupamento de termos;
- Classificação de documentos.

Funcionamento

- Terças-feiras: Laboratório. Quintas-feiras: Sala de aula;
- **As aulas práticas foram feitas para serem seguidas!**
- Trabalhos da disciplina desenvolvidos em Python;
- Algumas aulas introdutórias sobre Python. Material para aprofundamento disponível:
 - Livro, videoaulas, notebooks

Funcionamento

- **Lema das aulas:**
 - Não tenha medo de perguntar! A pergunta idiota é aquela que você não faz!
- **Relacionamentos são baseados na confiança!**
 - Quando há bom comportamento, interesse, dedicação e honestidade nas avaliações e trabalhos, os alunos ganham confiança e apreço, e tudo corre melhor!

Ferramentas

- Microsoft Teams
- Python 3:
 - SpaCy
 - Numpy / Matplotlib
 - BeautifulSoup
- Google Colab