



## FICHA DE COMPONENTE CURRICULAR

<b>CÓDIGO:</b> GBC214	<b>COMPONENTE CURRICULAR:</b> RECUPERAÇÃO DA INFORMAÇÃO	
<b>UNIDADE ACADÊMICA OFERTANTE:</b> FACULDADE DE COMPUTAÇÃO		<b>SIGLA:</b> FACOM
<b>CH TOTAL TEÓRICA:</b> 60 horas	<b>CH TOTAL PRÁTICA:</b> 00 horas	<b>CH TOTAL:</b> 60 horas

### 1. OBJETIVOS

Capacitar o aluno a propor soluções para o problema de, a partir de uma consulta formulada por um usuário e de uma coleção documentos não estruturados ou semiestruturados, por exemplo base de dados em XML, textuais ou de imagens digitais, recuperar informações nos documentos da coleção que melhor atendam às necessidades do usuário. Além disso, capacitar o aluno a avaliar a qualidade de sistemas de recuperação de informação por meio de medidas propostas na literatura.

### 2. EMENTA

Conceitos de documento, palavra e termo. Indexação de documentos: extração de termos, *stopwords*, *stemming*, criação de índices. Thesauri. Modelos de processamento de consultas. Avaliação de Sistemas de RI. RI em Documentos semiestruturados, multimídia e documentos na Web. Extração da informação. Classificação de documentos.

### 3. PROGRAMA

1. Introdução à Recuperação de Informação e modelo booleano
2. Dicionário e lista de *postings* : conceitos de documento, palavra e termo.
3. Indexação de documentos: termos, stopwords, stemming, Thesauri.
4. Compressão de índices 5. Peso de termos.
6. Modelo Vetorial
7. Avaliação de sistemas de recuperação de informação
8. Realimentação de relevantes e expansão de consultas
9. Recuperação em documentos semiestruturados (XML)
10. Modelo Probabilístico
11. Classificação de documentos
12. Agrupamento de documentos
13. Redução de dimensionalidade

14. Web: busca, *crawling*, indexação, análise de links
15. Extração da informação
16. Introdução à Recuperação de Imagens baseada em conteúdo

#### 4. BIBLIOGRAFIA BÁSICA

1. MANNING, C.; RAGHAVAN, P.; SCHÜTZE, H. **An introduction to information retrieval**. Cambridge: Cambridge University Press, 2009.
2. BAEZA-YATES, R.; RIBEIRO NETO, B. **Modern information retrieval**. New York: ACM Press, c1999.
3. BUTTCHER, S.; CLARKE, C. L. A.; CORMACK, G. V. **Information retrieval: implementing and evaluating search engines**. Cambridge: MIT Press, 2010.

#### 5. BIBLIOGRAFIA COMPLEMENTAR

1. FRAKES, W. B.; BAEZA-YATES, R. (ed.). **Information retrieval & data structures**. Nova Jersey: Prentice Hall, 1992.
2. CRESTANI, F.; PASI, G. (ed.). **Soft computing in information retrieval: techniques and applications**. Heidelberg: Physica, 2000.
3. MOENS, M. F. **Information extraction: algorithms and prospects in a retrieval context**. Dordrecht: Springer, c2006.
4. ZIVIANI, N. **Projeto de algoritmos com implementações em Pascal e C**. 2. ed. São Paulo: Cengage Learning, c2011.
5. BAEZA-YATES, R.; RIBEIRO NETO, B. **Recuperação de informação: conceitos e tecnologias das máquinas de busca**. Porto Alegre: Bookman, 2013. *E-book*. Disponível em: <https://www.sistemas.ufu.br/biblioteca-gateway/minhabiblioteca/9788582600498>. Acesso em: 10 jul. 2023.

#### 6. APROVAÇÃO

Maria Adriana Vidigal de Lima  
Coordenadora do Curso de Ciência da  
Computação

Maurício Cunha Escarpinati  
Diretor da Faculdade de  
Computação



Documento assinado eletronicamente por **Maria Adriana Vidigal de Lima, Coordenador(a)**, em 26/01/2024, às 15:45, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Mauricio Cunha Escarpinati, Diretor(a)**, em 19/02/2024, às 11:33, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [https://www.sei.ufu.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **5134794** e o código CRC **F416FDEE**.

---

**Referência:** Processo nº 23117.053855/2023-26

SEI nº 5134794