



UNIVERSIDADE FEDERAL DE OURO PRETO
DEPARTAMENTO DE COMPUTAÇÃO

PROGRAMA DE DISCIPLINA



Nome do Componente Curricular em português: Gerência de Dados na Web			Código: BCC450
Nome do Componente Curricular em inglês: Web Data Management			
Nome e sigla do departamento: Departamento de Computação (DECOM)			Unidade acadêmica: ICEB
Carga horária semestral: 60 horas	Carga horária semanal teórica: 4 horas/aula	Carga horária semanal prática: -	
Ementa: Caracterização dos dados da Web; modelagem de dados semiestruturados; linguagens de consulta; extração de dados da Web; integração de dados da web; aplicações.			
Conteúdo Programático: - Dados da Web - Características dos dados da Web - Web versus bancos de dados - Arquitetura das aplicações Web - Modelagem de dados semi-estruturados e XML - Características dos dados semi-estruturados - Modelos para representação de dados semi-estruturados - XML - Representação relacional de dados XML - Publicação de dados relacionais através de XML - Descoberta de estrutura de dados XML - Acessórios para gerência de dados XML - Compressão de dados XML - Linguagens de consulta - Linguagens de consulta para a Web - Linguagens de consulta para dados semi-estruturados - Linguagens de consulta e manipulação de dados XML - Extração de dados de fontes da Web - Caracterização do problema - Principais abordagens - Ferramentas - Avaliação da qualidade das extrações - Integração de dados da Web - Caracterização do problema de integração - Abordagens e arquiteturas - Fusão de dados - Medidas de similaridade			

- Tópicos avançados
 - Processamento de consultas em documentos XML
 - Acesso a dados da Web oculta
 - Reconhecimento e extração de entidades
 - Web semântica
 - Web de dados
 - Coleta de páginas da Web
 - Bibliotecas digitais

Bibliografia Básica:

- ABITEBOUL, Serge; BUNEMAN, Peter; SUCIU, Dan. Gerenciando dados na WEB. Rio De Janeiro: Campus, 2000.
- GRAVES, Mark. Projeto de banco de dados com XML. São Paulo: Pearson Education, 2003.
- BAEZA-YATES, R; RIBEIRO-NETO, Berthier. Modern information retrieval. New York: Addison Wesley, 2011.
- BREITMAN, Karin Koogan. Web semântica: a internet do futuro. Rio de Janeiro: LTC, 2006.

Bibliografia Complementar:

- ANTONIOU, G; VAN HARMELEN, Frank. A semantic web primer. Cambridge: MIT Press 2008.
- LIU, Bing. Web data mining: exploring hyperlinks, contents, and usage data . Berlin, New York: Springer, 2007.
- KANTARDZIC, Mehmed. Data mining: concepts, models, methods, and algorithms. Hoboken, NJ: Wiley-Interscience IEEE Press, 2003.
- ELMASRI, Ramez; NAVATHE, Sham. Sistemas de banco de dados. 6. ed. São Paulo: Pearson Addison Wesley, 2011.
- GARCIA-MOLINA, Hector; ULLMAN, Jeffrey D.; WIDOM, Jennifer. Database systems: the complete book . 2. ed. Upper Saddle River: Pearson Prentice Hall, 2009.