


Tóp. de Banco de Dados III		
CENTRO UNIVERSITÁRIO CTC DEPARTAMENTO DE INFORMÁTICA		
		Ano e semestre 2022.1
INF2032	INF2032 – Tópicos em Banco de Dados III – “Criação e manutenção de grandes bases de conhecimento”	
	CARGA HORÁRIA TOTAL:	CRÉDITOS: 3
Pré-requisitos: Não há Prof. Antônio L. Furtado / Marco A. Casanova		

OBJETIVOS	Esta disciplina aborda a criação e manutenção de grandes bases de conhecimento para aplicações em Ciência de Dados.
EMENTA	Fundamentos e arquiteturas para bases de conhecimento; integração de conhecimentos de fontes diversas; descoberta e padronização da representação de entidades; descoberta e padronização da representação de atributos e relacionamentos; construção de esquemas abertos; avaliação da qualidade de bases de conhecimento; estudos de caso.
PROGRAMA	<ul style="list-style-type: none"> • Motivação: histórico; casos de uso; criação de bases de conhecimento (KBs) • Fundamentos e arquiteturas: representação de conhecimento; construção de KBs • Integração de conhecimentos de fontes diversas: alinhamento; alinhamento baseado em grafos • Descoberta e padronização da representação de entidades: uso de dicionários; métodos baseados em padrões; métodos baseados em aprendizado de máquina; <i>entity embedding</i> • Descoberta e padronização da representação de atributos e relacionamentos: métodos baseados em padrões e em regras; extração a partir de dados semi-estruturados; extração baseada em redes neurais; • Construção de esquemas abertos: extração de entidades; descoberta de propriedades; canonização • Avaliação da qualidade de bases de conhecimento: avaliação da qualidade em KBs; completude de KBs; regras e restrições; <i>graph embeddings</i> • Estudos de caso: YAGO; DBpedia; NELL; Wikidata; KBs industriais
AVALIAÇÃO	O desempenho do aluno será avaliado por seminários apresentados em sala.

BIBLIOGRAFIA PRINCIPAL	Gerhard Weikum, Xin Luna Dong, Simon Razniewski and Fabian Suchanek (2021), "Machine Knowledge: Creation and Curation of Comprehensive Knowledge Bases", Foundations and Trends® in Databases: Vol. 10, No. 2-4, pp 108–490. DOI: 10.1561/19000000064.
BIBLIOGRAFIA COMPLEMENTAR	(artigos selecionados)