


PERÍODO: 2019.1		
CENTRO UNIVERSITÁRIO CTC DEPARTAMENTO DE INFORMÁTICA		
INF2031	Tópicos de Banco de Dados II Prof. Marco Antonio Casanova	
	CARGA HORÁRIA TOTAL: 45	CRÉDITOS: 3
	Título do Tópico: Integração de Dados para Ciência de Dados Pré-requisito:	
OBJETIVOS	Esta disciplina aborda tecnologias para integração de grandes volumes de dados, com aplicações em Ciência de Dados.	
EMENTA	Integração de Dados; Ligação de Entidades; Extração de Dados; Fusão de Dados; Alinhamento de Esquemas; Uso de ontologias em Integração de Dados.	
PROGRAMA	<ul style="list-style-type: none"> • Introdução: O que é Integração de Dados; Aprendizado de Máquina para Integração de Dados; Integração de Dados em Ciência de Dados • Cenários de Integração de Dados: Convencional x Web; <i>Data Warehouse</i> x Integração Virtual • Processos para Integração de Dados: Ligação de Entidades; Extração de Dados; Fusão de Dados; Alinhamento de Esquemas • Ligação de Entidades: estratégias baseadas em regras; estratégias supervisionadas; estratégias baseadas <i>Deep Learning</i> • Extração de Dados: estratégias baseadas em regras; estratégias específicas para um domínio; estratégias probabilísticas; estratégias baseadas <i>Deep Learning</i> • Fusão de Dados: estratégias baseadas em regras; estratégias específicas para um domínio; estratégias probabilísticas; estratégias baseadas <i>Deep Learning</i> • Alinhamento de Esquemas: estratégias baseadas em Lógica de Descrição; estratégias para mapeamento semi-automático; <i>pay-as-you-go</i> 	

	<p><i>dataspaces</i>; estratégias baseadas em Lógicas e em <i>Deep Learning</i></p> <ul style="list-style-type: none">• Uso de ontologias em Integração de Dados
AVALIAÇÃO	O desempenho do aluno será avaliado por seminários apresentados em sala e uma monografia preparada para o curso.
BIBLIOGRAFIA PRINCIPAL	Artigos selecionados
BIBLIOGRAFIA COMPLEMENTAR	Artigos selecionados