

GERÊNCIA DE DADOS EM LARGA ESCALA

Tipo de Disciplina	Optativa	Carga Horária	67				
Pré-Requisitos	Engenharia de Dados						
Docente	Diego Ernesto Rosa Pessoa						
Distribuição da Carga Horária							
Teórica	30	Prática	37	EaD	0	Extensão	0
Ementa							
<p>Conceitos fundamentais, tecnologias e aplicações inovadoras voltadas para o processamento e análise de grandes volumes de dados (Big Data). Exploração dos princípios de escalabilidade e elasticidade aplicados à gerência de dados em larga escala. Análise das soluções tecnológicas mais recentes, incluindo Data Lakes, governança de dados, armazenamento distribuído e formatos de arquivo otimizados. Discussão de arquiteturas para pipelines de dados, processamento paralelo e distribuído, com foco em frameworks e ferramentas do estado da arte. Estratégias para otimização de performance e segurança em sistemas intensivos de dados. Estudo de casos práticos e aplicações em organizações e empresas, cobrindo o processamento de dados em tempo real e em batch, integração de dados de diferentes fontes, e soluções para recomendações, análise de logs e monitoramento de sistemas.</p>							
Bibliografia Básica							
<ol style="list-style-type: none">1. NELSON, Catherine. Engenharia de Software para Cientistas de Dados: De Notebooks a Sistemas Escaláveis. Novatec, 2024. ISBN: 978-85752291702. REIS, Joe; HOUSLEY, Matt. Fundamentos de Engenharia de Dados: Projete e Construa Sistemas de Dados Robustos. Novatec, 2023. ISBN: 978-85752287603. SERRA, James. Decifrando Arquiteturas de Dados: Escolhendo entre data warehouse moderno, data fabric, data lakehouse e data mesh. Novatec, 2024. ISBN: 978-8575229217							
Bibliografia Complementar							
<ol style="list-style-type: none">1. AMARAL, Fernando. Aprenda mineração de dados ; teoria e prática / Fernando Amaral. Rio de Janeiro : Alta Books, 2016.2. BENGFORT, Benjamin. Analítica de dados com Hadoop : introdução para cientistas de dados. São Paulo: Novatec, 2016.3. CHEN, Daniel. Análise de dados com Python e Pandas / Daniel Y. Chen.- São Paulo: Novatec, 2018.4. MACHADO, Felipe Nery Rodrigues. Big data : o futuro dos dados e aplicações. São Paulo: Érica, 2018.5. RAMOS, Atos. Infraestrutura BIG DATA com OpenSource. Rio de Janeiro: Ciência Moderna, 2015.							
Bibliografia Suplementar (Periódicos)							
Publicações indexadas no Portal de Periódicos da CAPES (www.periodicos.capes.gov.br), no Repositório Institucional do IFPB (repositorio.ifpb.edu.br) e na Editora do IFPB (editora.ifpb.edu.br).							
Observações							