

CIN 7929 – PLANO DE ENSINO (2020.1)

1. IDENTIFICAÇÃO

Disciplina: CIN7929 – Engenharia de Dados

Carga Horária: 36 h/a – 2 créditos (18 h/a Síncronas e 18 h/a Assíncronas)

Professor: Douglas Dyllon Jeronimo de Macedo – *douglas.macedo@ufsc.br*

Pré-requisito: Obrigatória. Sem pré-requisitos.

Horário: Quinta-feira das 10h10 às 11h50

Local: Ensino remoto, devido a crise do COVID-19.

2. EMENTA

Armazenamento de dados. Bancos de Dados Distribuídos. Computação em Nuvem. Desenvolvimento de Arquiteturas.

3. OBJETIVOS

3.1 Objetivo Geral

Conhecer os fundamentos sobre engenharia de dados, assim como o desenvolvimento de arquiteturas de softwares robustas, flexíveis e adaptáveis.

3.2 Objetivos Específicos:

3.2.1. Conhecer os tipos de dados.

3.2.2. Conhecer os fundamentos de armazenamento de dados.

3.2.3. Estudar os fundamentos de sistemas distribuídos.

3.2.4. Entender os principais métodos e ferramentas para engenharia de dados.

4. CONTEÚDO PROGRAMÁTICO

4.1 Tipos de Dados

4.1.1 Dados estruturados

4.1.2 Dados semiestruturados

4.1.3 Dados não-estruturados

4.2 Armazenamento dos Dados

4.2.1 Modelos de armazenamento

4.2.2 SQL, NoSQL e NewSQL

4.2.3 Ferramentas para armazenamento de dados.

4.3 Sistemas Distribuídos

4.3.1 Computação Distribuída

4.3.1.1 Agregados Computacionais.

4.3.1.2 Grades Computacionais

4.3.1.3 Nuvens Computacionais

4.3.1.4 Outros modelos.

4.3.2 Bancos de Dados Distribuídos

4.4 Desenvolvimento de Arquiteturas

4.4.1 Aspectos gerais.

4.4.2 Definição de requisitos.

4.4.3 Ferramentas para Apoio e Desenvolvimento de Arquiteturas.

4.4.4 Segurança de Dados

5. METODOLOGIA

- Todas as aulas serão ministradas na modalidade não presencial, utilizando-se de atividades pedagógicas síncronas e assíncronas.
- Teremos como apoio da disciplina o **Moodle** institucional da UFSC (<http://moodle.ufsc.br>) e o mesmo será utilizado como instrumento de suporte da disciplina, devendo portanto, ser acessado todas as semanas pelos alunos.
- A organização dos conteúdos se dará através de tópicos, sendo que cada um terá um fórum de discussão (utilizando a funcionalidade já disponível no AVA Moodle) para que os alunos possam debater suas opiniões e dúvidas.
- Os encontros síncronos têm como objetivo a discussão de conteúdos e resolução de dúvidas. Estes encontros serão nos horários pré-definidos da disciplina e terão duração média de 1h, podendo ser prorrogados ao limite do tempo da disciplina de acordo com a demanda dos estudantes.
- As atividades síncronas serão realizadas utilizando as plataformas Mconf, Google Meet e/ou Jitsi, com a definição das salas virtuais previamente divulgada aos alunos via Moodle. Recomenda-se fortemente estudar previamente o material disponibilizado.
- As atividades assíncronas prevêm a distribuição dos conteúdos através de textos em PDF, videoaulas curtas, referências a outros materiais disponíveis online, questionários, fóruns de discussão e exercícios.

6. AVALIAÇÃO

Conforme a Resolução do Conselho Universitário 017/Cun/97, de 30/09/97:

- A frequência mínima obrigatória é de 75% das aulas.
 - O aluno deverá entrar no Moodle e marcar sua presença.

- Em caso de falta em dia de avaliação (prova, apresentação de trabalho ou mini-seminário), encaminhar justificativa e pedido formal à Chefia do Departamento de Ciência da Informação, no prazo de três (3) dias úteis.
- O aluno que obtiver nota final entre 3,0 e 5,5 e frequência suficiente poderá, ao final do semestre, realizar uma prova de recuperação de todo o conteúdo.
- A nota final será calculada a partir da média simples entre a nota final obtida durante o semestre e a nota obtida na prova de recuperação.

Avaliações:

- Serão realizados diversos exercícios durante o semestre (E1).
- Será realizada uma atividade (A1).
- Será realizado um trabalho final (T1).
- A média final (MF) será calculada conforme a seguinte fórmula:

$$\mathbf{MF = (E1 + A1 + T1) / 3}$$

Entregas:

- E1: todos os exercícios devem ser entregues via Moodle.
- A1: o endereço do blog/site deverá ser postado no Moodle.
- T1: um vídeo deverá ser elaborado pela equipe e postado em alguma plataforma online. A entrega será feita via Moodle, através do link do vídeo.

7. BIBLIOGRAFIA

7.1 Bibliografia básica

COULOURIS, George F.; DOLLIMORE, Jean; KINDBERG, Tim. Sistemas distribuídos: conceitos e projeto. 4. ed. Porto Alegre: Bookman, 2007. viii, 784 p. ISBN 9788560031498. (4 – BU/UFSC)

DATE, C. J. Introdução a sistemas de bancos de dados. Rio de Janeiro: Campus, 2004. 865p. ISBN 8535212736. (2 – BU/UFSC)

HOWS, D.; MEMBREY, P.; PLUGGE, E. Introdução ao MongoDB. São Paulo, Novatec, 2015. (0 – BU/UFSC)

PROVOST, Foster; FAWCETT, Tom. Data science for business. Sebastopol: O'Reilly, c2013. xxi, 386 p. ISBN 9781449361327. (5 – BU/UFSC)

Apoio de Material Online:

MACEDO, Douglas Dyllon Jeronimo de. Um modelo distribuído de armazenamento hierárquico de conhecimento médico. 2014. 1 v. Tese (Doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Florianópolis, 2013. Disponível em: <<http://www.bu.ufsc.br/teses/PEGC0324-T.pdf>

INACIO, Eduardo Camilo. Caracterização e modelagem multivariada do desempenho de sistemas de arquivos paralelos. 2015. 96 p. Dissertação (Mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Ciência da Computação, Florianópolis, 2015. Disponível em: <<http://www.bu.ufsc.br/teses/PGCC1026-D.pdf>

POFFO, João Paulo. Projeto lógico de bancos de dados NOSQL colunares a partir de esquemas conceituais entidade-relacionamento estendido (EER). 2016. 85 p. Dissertação (Mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Ciência da Computação, Florianópolis, 2016 Disponível em: <<http://www.bu.ufsc.br/teses/PGCC1070-D.pdf>

7.2 Bibliografia complementar

ATZENI, Paolo. Database systems: concepts, languages & architectures. London: McGraw Hill, c2000. xxiii, 612p. ISBN 007235872.

HURWITZ J. et al. Big Data for Dummies. New Jersey (USA): John Wiley & Sons, Inc., 2013. ISBN 978-1-118-50422-2.

MAYER-SCHÖNBERGER, Viktor; CUKIER, Kenneth. Big data: a revolution that will transform how we live, work, and think. Boston: Houghton Mifflin Company, c2013. 252 p. ISBN 9780544227750.

Özsu, M. Tamer, and Patrick Valduriez. Principles of distributed database systems. Springer Science & Business Media, 2011.

8. CRONOGRAMA ¹				
#	Data	Tipo	Conteúdo	C.H.
1	05/03	Presencial	Apresentação da disciplina e do plano de ensino.	2
2	12/03	Presencial	Engenharia e a Ciência de Dados	2
3	03/09	Síncrono	Webconferência Realinhamento e atualização da disciplina. Explicação da Metodologia Apresentação: Parte 1 - Engenharia e a Ciência de Dados	2
4	10/09	Assíncrono	Parte 1 - Engenharia e a Ciência de Dados Leitura dos Slides Desenvolvimento - Exercício 1 Leitura do Artigo 1: Ciência de Dados	2
5	17/09	Assíncrono	Parte 1 - Engenharia e a Ciência de Dados Leitura do Artigo 2: Data Science Entrega do Exercício 1	2
6	24/09	Síncrono	Webconferência	2

			Apresentação: Parte 2 - Sistemas Distribuídos	
7	01/10	Síncrono	Webconferência Apresentação: Parte 3 - Computação Distribuída	2
8	08/10	Assíncrono	Parte 3 - Computação Distribuída Exercício 1 e 2	2
9	15/10	Síncrono	Webconferência Parte 4 - Armazenamento de Dados Explicação e Definição da Atividade (A1) - Blog ou Site.	2
10	22/10	Assíncrono	Parte 4 - Armazenamento de Dados Desenvolvimento da Atividade (A1)	2
11	29/10	Assíncrono	Parte 4 - Armazenamento de Dados Desenvolvimento da Atividade (A1)	2
12	05/11	Assíncrono	Parte 4 - Armazenamento de Dados Entrega da Atividade (A1)	2
13	12/11	Síncrono	Webconferência Parte 5 - Desenvolvimento de Arquiteturas Explicação e Definição dos Trabalhos Finais (T1)	2
14	19/11	Assíncrono	Parte 5 - Desenvolvimento de Arquiteturas Explicação e Definição dos Trabalhos Finais (T1)	2
15	24/11	Assíncrono	Parte 5 - Desenvolvimento de Arquiteturas Desenvolvimento do Trabalho Final (T1)	2
16	03/12	Síncrono	Webconferência Parte 5 - Desenvolvimento de Arquiteturas Desenvolvimento do Trabalho Final (T1)	2
17	10/12	Assíncrono	Trabalho Final (T1) Entrega do Trabalho Final (T1) - Vídeo em plataforma online. Todos os alunos devem assistir todos os vídeos dos colegas. Fórum - Perguntas e Respostas	2
18	17/12	Síncrono	Recuperação Final.	2
TOTAL				36

NOTA IMPORTANTE

- Todas as aulas estão protegidas pelo direito autoral. Baixar, reproduzir, compartilhar, comunicar ao público, transcrever, transmitir, entre outros, o conteúdo das aulas ou de qualquer material didático-pedagógico só é possível com prévia autorização.
- Respeite a privacidade e os direitos de imagem tanto dos docentes quanto dos colegas. Não compartilhe *prints*, fotos, etc., sem a permissão explícita de todos os participantes.
- O(a) estudante que desrespeitar esta determinação estará sujeito(a) a sanções disciplinares previstas no Capítulo VIII, Seção I, da Resolução 017/Cun/1997.