

DOSSIER DA DISCIPLINA

SDSD - SISTEMAS DE DADOS PARA SUPORTE À DECISÃO



LMCC (4º ano) + LESI (4º ano)
Docente responsável: Orlando Belo
Ano lectivo de 2003/2004 - 2º semestre

ÍNDICE

Apresentação	1
Identificação da disciplina	1
Historial	1
Objectivos	2
Equipa docente e horários	2
Turnos	3
Programa	4
Programa resumido	4
Programa detalhado	4
Avaliação	6
Regime de faltas	6
Métodos de avaliação	6
Elementos de avaliação	8
Classificações	8
Bibliografia	10
Bibliografia essencial	10
Bibliografia complementar	10
Material pedagógico	12
Referências Web	12
Ferramentas	13
Exames	13
Casos para Estudo	13
Artigos para Leitura	14
Aulas	14
Sumários	16

1 - APRESENTAÇÃO

1.1 - Identificação da disciplina

Disciplina	SDSD - Sistemas de Dados para Suporte à Decisão
Área científica	Data Warehousing
Instituição	Universidade do Minho
Conselho de Cursos	Escola de Engenharia
Departamento	Departamento de Informática
Cursos	LMCC , 4º ano (opcional - 7009U2) LESI , 4º ano (opcional - 5309V7)
Responsável	Orlando Belo
Escolaridade	2h(T) + 2h(TP) + 0h(P)
Unidades de Crédito	3.5
Regime	Semestral (2º semestre)
Ano lectivo	2003/2004

1.2 - Historial

A disciplina de Sistemas de Dados para Suporte à Decisão é uma disciplina semestral, proposta como Opção II, que está integrada no plano de estudos da Licenciatura em Engenharia de Sistemas e Informática e Licenciatura em Matemática e Ciências de Computação, no 4º ano, 2º Semestre, com um valor de 3.5 unidades de crédito. A escolaridade da disciplina está estruturada em 2/2/0, o que corresponde respectivamente à leccionação de 1 aula teórica (2 horas) e de 1 aula teórico/Prática (2 horas).

A disciplina foi leccionada pela primeira vez no ano lectivo de 1998/1999 com o objectivo de responder a uma necessidade emergente nessa altura no mercado de trabalho.

Como disciplina fundamental na área de sistemas de suporte à decisão, tem tido ao longo dos anos da sua leccionação particular ênfase nos domínios da modelação dimensional, extracção de dados em fontes de informação, optimização de interrogações (queries) em sistemas de dados de grande dimensão, desenho, implementação e exploração de data warehouses e sistemas de processamento analítico.

Nos últimos três anos de leccionação o número de alunos, das duas licenciaturas alvo da disciplina, tem vindo a aumentar gradualmente, a taxa de assiduidade às

aulas, tanto teóricas como teórico-práticas, tem-se mantido elevada (cerca de 90%).

1.3 - Objectivos

A disciplina de **Sistemas de Dados para Suporte à Decisão - Data Warehousing**, pretende transmitir aos alunos conhecimentos relacionados com:

- A área dos Sistemas de Suporte à Decisão (*Business Intelligence*), em termos gerais e com o projecto de Sistemas de *Data Warehousing*.
- A análise do desempenho dos sistemas de bases de dados convencionais no fornecimento de informação para suporte à decisão.
- A importância do desenvolvimento de sistemas de *Data Warehousing* na implementação de Sistemas de Suporte à Decisão.
- A migração de dados de sistemas operacionais para sistemas de *Data Warehousing* e para plataformas de processamento analítico.
- A análise, desenho, desenvolvimento e exploração de *DataMarts*.
- O desenvolvimento de sistemas e aplicação de técnicas de processamento analítico.

1.4 - Equipa docente e horários

Docentes



Orlando Belo (42835)

Responsável pela disciplina
Professor associado
Departamento de Informática
Escola de Engenharia
Universidade do Minho
Tel. : 253-604443/30
Fax : 253-604471
Gabinete : DI-3.24
URL : <http://www.di.uminho.pt/~omb>
Email : obelo@di.uminho.pt

Horários

Turno	Dia	Hora	Sala	Docente
T	5ª feira	11h-13h	DI-A1	Orlando Belo
TP	5ª feira	16h-18h	DI-1.09	Orlando Belo
atendimento	4ª feira	15h-17h	gabinete	Orlando Belo

1.5 - Turnos

Turno 1

Nome	T
Tipo	teórico
Cursos	LMCC + LESI
Docentes	Orlando Belo

Turno 2

Nome	TP
Tipo	teórico-prático
Cursos	LMCC + LESI
Docentes	Orlando Belo

Turno 3

Nome	atendimento
Tipo	teórico-prático
Cursos	LMCC + LESI
Docentes	Orlando Belo

2 - PROGRAMA

2.1 - Programa resumido

O programa da disciplina está organizado da seguinte forma:

1. **Sistemas de Suporte à Decisão**
2. **Data Warehouses**
3. **Processamento Analítico de Dados**

2.2 - Programa detalhado

1. Sistemas de Suporte à Decisão

- Evolução dos Sistemas de Suporte à Decisão.
- A necessidade de Sistemas de Suporte à Decisão.
- O ciclo de vida do suporte à decisão.
- Implementação de Sistemas de Suporte à Decisão.
- Processamento operacional .vs. analítico.
- Os Data Warehouses e os Sistemas de Suporte à Decisão.

2. Data Warehouses

- Ambiente e estrutura funcional de um Data Warehouse.
- O desenvolvimento incremental de um Data Warehouse.
- Granulosidade, particionamento e estruturação da Informação num Data Warehouse.
- Os componentes de um Data Warehouse.

- Desenho de um Data Warehouse.
- Modelos de dados para um Data Warehouse.
- Esquemas Conceptuais para um Data Warehouse.
- Extração, preparação, transformação e integração de dados.
- Integração de perfis de informação.
- Metadados.
- Data Warehouse departamentais ou Data Marts.
- Gestão e administração de Data Warehouses.
- Implementação de Data Warehouses em Plataformas Cliente/Servidor.
- Data Warehouses distribuídos.
- Optimização de queries em sistemas de dados de grande dimensão.
- Técnicas de scripting em processos de alimentação de sistemas de Data Warehousing.
- Web Data Warehouses.

3. Processamento Analítico de Dados

- Conceitos e terminologia base.
- Vantagens do processamento analítico.
- Modelos de dados multidimensionais.
- OLAP - On-Line Analytical Processing.
- Processamento analítico multidimensional .vs. multirelacional.
- Arquitecturas para processamento analítico de dados.

3 - AVALIAÇÃO

3.1 - Regime de faltas

O funcionamento da disciplina de SDSA está organizado em:

- **aulas teóricas (2h/sem)** : são utilizadas para a exposição, análise e debate de conceitos necessários à compreensão do tema da disciplina; o material leccionado nestas aulas é encadeado, pelo que se torna indispensável o seu regular acompanhamento.
A frequência às aulas teóricas é recomendada.
- **aulas teórico-práticas (2h/sem)** : têm por objectivo o desenvolvimento de competências intelectuais - aplicar e/ou aprofundar conceitos adquiridos nas aulas teóricas - e de competências técnicas - desenvolvimento de técnicas de aplicação dos conceitos à resolução de problemas.
A participação nas aulas teórico-práticas é recomendada, sendo obrigatória para os estudantes de 1ª inscrição; para os estudantes de 1ª inscrição existe controlo de presenças, para efeitos de cumprimento do RIAPA (com mais de 1/3 de faltas o aluno poderá não ser admitido a exame).

3.2 - Métodos de avaliação

Elementos de avaliação

A avaliação da aprendizagem da disciplina é de tipo não contínuo, envolvendo 2 elementos de avaliação:

i) Exame (ex)

A nota mínima admissível no exame é de 10 (dez) valores.

Os alunos que obtenham uma nota de 8 ou 9 em exame serão submetidos a uma prova oral.

ii) Trabalho Prático (tp)

A nota mínima admissível no trabalho prático é de 10 (dez) valores.

O trabalho prático deverá ser realizado em grupo. O número máximo de elementos por grupo de trabalho é de 5. Os relatórios finais dos trabalhos serão enviados por

correio electrónico até às 24 horas do dia da data de entrega. Porém podem ser definidas outras datas anteriores para verificação da evolução dos trabalhos de cada grupo.

As apresentações dos trabalhos serão escalonadas após a data de entrega do relatório final do trabalho e consistirão: numa apresentação em PowerPoint e numa demonstração do software desenvolvido.

O trabalho prático a realizar pelos alunos consistirá na análise, planeamento e projecto de um Sistema de Data Warehousing para um domínio de aplicação específico.

O trabalho prático será realizado por etapas, cada uma delas correspondendo a um etapa específica do processo de análise, projecto e implementação do sistema referido. No final de cada uma das etapas, cada um dos grupos de trabalho deverá entregar, por email e em formato pdf, o relatório da etapa realizada. Em cada uma das etapas, será fornecido um template específico para a realização do respectivo relatório. No início do projecto cada grupo de trabalho deverá preencher e entregar a respectiva ficha de projecto requerida pelo responsável da disciplina.

Os prazos de entrega para os relatórios intermédio e final são os seguintes:

- 1ª Etapa : Arranque do projecto - 22/Março/2004
- 2ª Etapa : Planeamento do projecto, definição dos requisitos de negócio - 5/Abril/2004
- 3ª Etapa : Modelação dimensional, desenho físico, desenvolvimento da área de retenção - 26/Abril/2004
- 4ª Etapa : Projecto técnico da arquitectura do sistema, selecção e instalação de produtos - 17/Maio/2004
- 5ª Etapa : Análise e avaliação do trabalho realizado - 24/Maio/2004
- Etapa final : Relatório final do projecto - 5/Junho/2004

Fórmula de avaliação

A classificação final **nf** será dada pela fórmula :

$$nf = 50 \% * ex + 50 \% * tp$$

Resumo

A realização das provas escritas nas diferentes épocas de exame, a admissão às

provas orais, bem como os demais elementos de avaliação, decorrem de acordo com o estabelecido pelo RIAPA, aprovado pelo Despacho RT-20/2000 e pelo Despacho RT-43/2000.

3.3 - Elementos de avaliação

Nesta secção serão disponibilizados ao longo do período lectivo os enunciados dos elementos de avaliação que compõem o sistema de avaliação da disciplina, assim como material de apoio anexo aos enunciados.

Exame

- Enunciado do exame da época normal (1ª chamada) [[pdf](#)]
- Enunciado do exame da época normal (2ª chamada) [[pdf](#)]
- Enunciado do exame da época de recurso [[pdf](#)]

Trabalho Prático

- Enunciado do trabalho prático [[pdf](#)]
- Ficha de projecto [[pdf](#)] [[doc](#)]
- Template para a 1ª etapa [[doc](#)]
- Template para a 2ª etapa [[doc](#)]
- Template para para a 3ª etapa [[doc](#)]
- Template para a 4ª etapa [[doc](#)]
- Template para a 5ª etapa [[doc](#)]
- Template para o relatório final [[doc](#)]

3.4 - Classificações

Nesta secção serão dadas a conhecer as classificações de todos os elementos de avaliação que compõem o sistema de avaliação da disciplina.

- Pauta da época normal (1ª chamada) [[pdf](#)]
(04-07-10)
- Pauta da época normal (2ª chamada) [[pdf](#)]
(04-07-10)
- Pauta da época de recurso [[pdf](#)]
(04-07-30)

4 - BIBLIOGRAFIA

4.1 - Bibliografia essencial

- [KRRT98] *Kimball, R., Reeves, L., Ross, M., Thornthwait, W., The Data Warehouse Lifecycle Toolkit - Expert Methods for Designing, Developing, and Deploying Data Warehouses*, John Wiley & Sons, 1998.
ISBN : 0-471-25547-5

Na realização do trabalho prático, cada grupo de projecto deverá seguir a metodologia apresentada e discutida nesta obra.

4.2 - Bibliografia complementar

- [DD97] *Date, C.J., Darwen, H., A Guide to the SQL Standard*, IV Edição, Addison-Wesley Inc, 1997.
- [In96] *Inmon, W.H., Building the Data Warehouse*, John Wiley & Sons, 1996.
URL : http://www3.baylor.edu/~Xin_Wang/pdf/datawarehouse.pdf
- [Kim96] *Kimball, R., Data Warehouse Toolkit : Practical Techniques for Building Dimensional Data Warehouses*, John Wiley & Sons, John Wiley & Sons, 1996.
- [BS97] *Berson, A., Smith, S.J., Data Warehousing, Data Mining, & OLAP*, McGraw-Hill, 1997.
- [KM00] *Kimball, R., Merz, R., Data Webhouse Toolkit : Building the Web-Enabled Data Warehouse*, John Wiley, 2000.
- [Teo94] *Teorey, T., Database Modeling and Design: The Fundamental Principles*, II Edição, Morgan Kaufmann, 1994.
- [CB01] *Connolly, T., Begg, C., Database Systems - A Practical Approach*

to Design, Implementation, and Management , III Edição ,
Addison-Wesley , 2001.

[IWG97]

Inmon, W.H., Welch, J.D., Glassey, K.L., **Managing the Data Warehouse** , John Wiley & Sons , 1997.

[CAABTV99]

Corey, M., Abbey, M., Abramson, I., Barnes, L., Taub, B., Venkitachalam, R., **SQL Server / Data Warehousing** ,
Osborne/McGraw-Hill , 1999.

5 - MATERIAL PEDAGÓGICO

5.1 - Referências Web

- <http://www.oracle.com/>
Oracle Corporation
Referências a todas as famílias de produtos da Oracle. Acesso à página do sistema de gestão de bases de dados Oracle9i através do endereço www.oracle.com/ip/deploy/database/9i e Oracle8i através do endereço www.oracle.com/ip/deploy/database/8i.
- <http://www.ibm.com/>
IBM Corporation
Referências a todas as famílias de produtos da IBM. Acesso à página do sistema de gestão de bases de dados DB2 através do endereço www-4.ibm.com/software/data/db2.
- <http://www.microsoft.com/>
Microsoft Corporation
Referências a todas as famílias de produtos da Microsoft. Acesso à página do sistema de gestão de bases de dados SQL Server 2000 através do endereço www.microsoft.com/sql e Access 2000 através do endereço www.microsoft.com/office/access.
- <http://www.informix.com/>
IBM Corporation
Site actualmente da IBM que acolhe a família de produtos da "antiga" Informix Corporation. O endereço redirecciona o utilizador para o endereço www-4.ibm.com/software/data/informix localizado já no site de IBM.
- <http://www.sybase.com/>
Sybase Corporation
Referências a todas as famílias de produtos da Sybase. De realçar as referências relativas a Servidores de bases de dados, servidores aplicativos, business intelligence e ferramentas para modelação de esquemas de dados.
- <http://www.ingres.com/>
Computer Associates
Ingres II Sistema de gestão de bases de dados e de recursos aplicativos.
- <http://www.embarcadero.com/>
Embarcadero Technologies
Ferramentas de desenho, optimização de desempenho, debugging, e

recuperação de sistemas de bases de dados.

- <http://www.cai.com/products/alm/Erwin.htm>
Computer Associates-ERwin
Ferramentas de desenho e modelação de esquemas para sistemas de bases de dados.
- <http://citeseer.nj.nec.com/cs>
The NECI Scientific Literature Digital Librar
Pesquisador de referências bibliográficas. Possibilita o acesso a um conjunto muito vasto de publicações.

5.2 - Ferramentas

- Microsoft MS SQL Server 2000
<http://www.microsoft.com/sql>
- Sybase Power Designer
<http://www.sybase.com/>
- Microsoft Visio 2000
<http://www.microsoft.com/>
- Embarcadero Studio
<http://www.embarcadero.com/>
- Computer Associates ERWin
<http://www.cai.com/products/alm/Erwin.htm>

5.3 - Exames

- 2002-2003 : [1ª Chamada](#), [2ª Chamada](#), [Recurso](#).
- 1999-2000 : [1ª Chamada](#), [2ª Chamada](#), [Recurso](#).

5.4 - Casos para Estudo

- [Caso 01](#) - "Gestão de Custos de Comunicações".
- [Caso 02](#) - "Centros de Comercialização de Produtos".

- [Caso 03](#) - "Uma Empresa de Distribuição de Água".
- [Caso 04](#) - "Um Data Mart para a 'Pubs'".
- [Caso 05](#) - "O Sistema de Suporte à Decisão da Portugalis".

5.5 - Artigos para Leitura

- J. Widom, ["Research Problems in Data Warehousing"](#). Proceedings of the 4th Int'l Conference on Information and Knowledge Management (CIKM), November 1995.
- D. Calvanese, G. De Giacomo, M. Lenzerini, D. Nardi, and R. Rosati. [Source integration in data warehousing](#). Technical Report DWQ-UNIROMA002, DWQ Consortium, Oct. 1997.
- W. Labio, D. Quass, B. Adelberg, ["Physical Database Design for Data Warehouses"](#), In Proceedings of the International Conference of Data Engineering.
- M. Chen, J. Han, and P. Yu. ["Data mining: An overview from database perspective"](#). IEEE Transactions on Knowledge and Data Eng., 8(6):866--883, December 1996.
- Hammer J., Garcia-Molina H., Widom J., Labio W, and Zhuge Y.: [The Stanford Data Warehousing Project](#). In IEEE Data Engineering Bulletin, 18(2), pp. 41--48, June 1995.
- M. Bouzeghoub, F. Fabret, M. Matulovic-Broqué, ["Modeling Data Warehouse Refreshment Process as a Workflow Application"](#), Proceedings of the International Workshop on Design and Management of Data Warehouses (DMDW'99), Heidelberg, Germany, 14. - 15. 6. 1999.
- A. Gutiérrez, A. Marotta, ["An Overview of Data Warehouse Design Approaches"](#), Apple, 2000.

5.6 - Aulas

i) Conjunto de slides das aulas teóricas

- [Sistemas de Data Warehousing](#) (V2.04), Orlando Belo, Departamento de

Informática, Universidade do Minho, 2004.

- [Metadados em Sistemas de Data Warehousing](#) (V2.04), Orlando Belo, Departamento de Informática, Universidade do Minho, 2003.

ii) Ficheiros das aulas práticas

– Parque de Cereais

1. Portões de Acesso

- Porta A - Ficheiro Texto [[txt](#)]
- Porta B - Ficheiro MS-Excel [[xls](#)]
- Porta C - Ficheiro MS-Access [[mdb](#)]
- Porta D - Ficheiro Texto [[txt](#)]

2. Base de Dados do Sistema Principal

Base de Dados ParqueCereais - Cópia de Segurança SQL2K [[bk](#)]

3. Sistemas de Transformação de Dados

Alimentação da base de dados ParqueCereais [[dts](#)]

6 - SUMÁRIOS

Semana 1

Aula nº 1

Tipo	T
Data	2003-02-26
Hora	11h-13h
Sumário	<hr/> <p>Apresentação da disciplina - objectivos, programa, bibliografia e critérios de avaliação. Breve apresentação da área dos Sistemas de Data Warehousing.</p> <hr/>

Aula nº 2

Tipo	TP
Data	2003-02-26
Hora	16h-18h
Sumário	<hr/> <p>Apresentação da componente teórico-prática da disciplina.</p> <hr/>

Semana 2

Aula nº 3

Tipo	T
Data	2003-03-04
Hora	11h-13h
Sumário	<hr/> <p>Sistemas de Data Warehousing - contextualização, objectivos, sistemas de suporte à decisão, cenários de gestão, exploração de dados e tomada de decisão; intervenientes e recursos envolvidos; análise e discussão de problemas sobre qualidade de informação.</p> <hr/>

Aula nº 4

Tipo TP
Data 2003-03-04
Hora 16h-18h
Sumário

Apresentação e demonstração das ferramentas de trabalho a utilizar na componente prática da disciplina.
 Análise de sistemas operacionais, análise de esquemas físicos de bases de dados; descrição, manipulação e controlo de esquemas conceptuais de fontes de informação relacionais; modelação de processos de transferência de dados; e definição de perfis de utilização e suas credenciais.

Semana 3

Aula nº 5

Tipo T
Data 2003-03-11
Hora 11h-13h
Sumário

Sistemas de Data Warehousing - fundamentação e bases para um sistema de data warehousing; arquitecturas naturalmente envolventes; implicações da qualidade de dados dos sistemas operacionais nos sistemas de suporte à decisão.
 Identificação e caracterização de problemas na satisfação de interrogações dos agentes de decisão em tempo útil.
 Justificação, passo a passo, da necessidade de um sistema de data warehousing.

Aula nº 6

Tipo TP
Data 2003-03-11
Hora 16h-18h
Sumário

Análise da qualidade de dados de um sistema operacional.
 Detecção, análise e caracterização de situações anómalas em fontes de dados relacionais e análise das suas implicações em

sistemas de suporte à decisão.
Estudo de um caso de aplicação prática.

Semana 4

Aula nº 7

Tipo	T
Data	2003-03-18
Hora	11h-13h
Sumário	

Apresentação do trabalho prático e planeamento das suas etapas principais.
Sistemas de Data Warehousing - Modelo de organização básico, características e funcionalidades principais; vantagens e problemas a considerar na implementação deste tipo de sistemas; requisitos operacionais; arquitectura funcional e seus principais componentes; data warehouses e data marts; fluxos de dados.

Aula nº 8

Tipo	TP
Data	2003-03-18
Hora	16h-18h
Sumário	

Análise e Modelação de Sistemas de Dados - análise, modelação e desenvolvimento de esquemas relacionais para o caso de aplicação real; planeamento e definição de interrogações; revisão de conceitos, terminologia e sistemas de modelação.
Introdução à Modelação Dimensional - conceitos básicos; características dos esquemas dimensionais - dimensões, medidas e esquemas estrela.
