

DOSSIER DA DISCIPLINA

PED - PROCESSAMENTO ESTRUTURADO DE DOCUMENTOS



Mestrado em Informática (1º ano) + Curso de Especialização em Informática (1º ano)

Docente responsável: José Carlos Ramalho

Ano lectivo de 2004/2005 - 2º semestre

ÍNDICE

Apresentação	1
Identificação da disciplina	1
Historial	1
Objectivos	1
Equipa docente e horários	2
Turnos	2
Programa	3
Programa resumido	3
Programa detalhado	3
Avaliação	7
Regime de faltas	7
Métodos de avaliação	7
Elementos de avaliação	8
Classificações	8
Bibliografia	9
Bibliografia essencial	9
Bibliografia complementar	9
Material pedagógico	10
Referências Web:	10
Sumários	11

1 - APRESENTAÇÃO

1.1 - Identificação da disciplina

Disciplina	PED - Processamento Estruturado de Documentos
Área científica	
Instituição	Universidade do Minho
Conselho de Cursos	Escola de Engenharia
Departamento	Departamento de Informática
Cursos	Mestrado em Informática , 1º ano (opcional -) Curso de Especialização em Informática , 1º ano (opcional -)
Responsável	José Carlos Ramalho
Escolaridade	2h(T)
Unidades de Crédito	1.5
Regime	Semestral (2º semestre)
Ano lectivo	2004/2005

1.2 - Historial

Esta é a quarta instância da disciplina na pós-graduação. Devido à elevada procura que tem tido, tem sido oferecida nos mesmos moldes mas com um programa actualizado ano a ano.

1.3 - Objectivos

Apesar de na definição da disciplina estar estipulado que esta terá uma aula teórica de 2 horas, o que acontece na realidade é que, cerca de 95% da disciplina, é leccionada em regime teórico-prático, ou seja, os conceitos são introduzidos com casos práticos que os alunos experimentam nas aulas.

Para a frequência desta disciplina não são necessários requisitos especiais. No entanto, os alunos com experiência em tecnologias Web e/ou programação declarativa sentirão que a disciplina é bastante acessível enquanto alunos com experiências adquiridas diferentes terão de trabalhar um pouco mais para atingir os mesmos resultados.

1.4 - Equipa docente e horários

Docentes

José Carlos Ramalho (406038)
Responsável pela disciplina
Professor auxiliar
Informática

Horários

1.5 - Turnos

Turno 1

Nome	T1
Tipo	teórico
Cursos	Mestrado em Informática + Curso de Especialização em Informática
Docentes	José Carlos Ramalho

2 - PROGRAMA

2.1 - Programa resumido

O programa da disciplina está organizado da seguinte forma:

1. Introdução
2. Documentação Estruturada e Anotação
3. Documentos XML: estrutura e conceitos
4. Ciclo de vida documental
5. Desenvolvimento de DTDs e Schemas
6. XSL: estrutura e conceitos
7. Processamento imperativo de documentos
8. XML: últimos desenvolvimentos
9. XML e o intercâmbio da informação
10. Algumas aplicações XML

2.2 - Programa detalhado

1. Introdução

Um pouco de história: SGML, HTML, XML, SML.

As origens destas normas. Como é que elas se inter-relacionam.

2. Documentação Estruturada e Anotação

O conceito de documento estruturado.

Os vários tipos de anotação e sua evolução nos últimos tempos.

As vantagens e as desvantagens da Anotação Descritiva.

Como é composta uma linguagem de anotação e qual a sua aplicação.

O conceito de meta-linguagem de anotação: o SGML e o XML.

3. Documentos XML: estrutura e conceitos

As componentes dum documento XML: a declaração, o DTD e a instância.

Documento válido versus documento bem-estruturado.

Elementos estruturais dum documento XML: elementos, atributos e entidades.

Alguns exemplos de documentos XML: o soneto, os sumários, algumas obras literárias e informação importada de bases de dados.

4. Ciclo de vida documental

Técnicas de Análise documental: diagramas, elm-trees.

Edição de documentos XML: utilização de editores estruturados.

Validação de documentos: integrada em editores ou outras ferramentas e na linha de comando.

Introdução à transformação e formatação de documentos.

Como armazenar documentos estruturados: XML e bases de dados.

5. Desenvolvimento de DTDs e Schemas

Construção de DTDs e sua utilização no ciclo de vida documental.

Princípios básicos a serem observados no desenvolvimento de DTDs. Algumas técnicas.

Introdução aos Schemas: DTDs versus Schemas.

Comparação de Schemas com gramáticas independentes de contexto.

Desenvolvimento de Schemas para alguns casos de estudo.

6. XSL: estrutura e conceitos

O modelo de dados subjacente a um documento XML.

Transformação de documentos como transformação de árvores abstractas.

O XSL e a transformação declarativa de documentos XML.

Estrutura duma especificação em XSL e respectivo modelo de processamento.

Os vários operadores e funções do XSL.

O XSLFO e a geração de formatos para distribuição a partir do XML.

7. Processamento imperativo de documentos

Utilização de linguagens de programação imperativas para transformar documentos XML.

Os dois modelos de programação: "tree-driven" (DOM), "event-driven" (SAX).

Alguns exemplos com aplicação: XML::DT, Omnimark.

8. XML: últimos desenvolvimentos

Algumas normas que irão interagir com o XML: XPointer, XLink, NameSpaces, XPipe, MDTs.

9. XML e o intercâmbio da informação

Introdução ao XML.

Alguns exemplos de aplicação.

10. Algumas aplicações XML

- EAD: Encoded Archive Description.
- DocBook: Scientific Documentation.
- HL7: Medical Records.
-

MathML: Mathematical Markup Language.

- XCSL: XML Constraint Specification Language.
- XTM: XML Topic Maps.
- WSDL: XML Web Services.
- SVG: Standard Vector Graphics.

3 - AVALIAÇÃO

3.1 - Regime de faltas

As faltas não contarão para efeitos de avaliação.

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 (componente-teórica)**

O aluno poderá ser avaliado por um exame convencional.

ii) **Projecto (componente-prática)**

Alternativamente, o aluno poderá ser avaliado através do desenvolvimento de um projecto.

O projecto e os parâmetros da avaliação prática são descritos num documento à parte que se inclui a seguir.

Fórmula de avaliação

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

$$\text{Max}(\textbf{nf} = 100 \% * \textbf{componente-teórica} , 100 \% * \textbf{componente-prática})$$

Resumo

O algoritmo para cálculo da nota final é o máximo dos elementos realizados pelo aluno.

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.

- (ainda não estão disponíveis os enunciados).

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.

- (ainda não existem classificações a disponibilizar).

4 - BIBLIOGRAFIA

4.1 - Bibliografia essencial

- [RH02] *José Carlos Ramalho, Pedro Rangel Henriques, XML e XSL: da teoria à prática*, 1ª Edição, FCA, Novembro de 2002.
ISBN :
Este livro foi escrito para suportar cursos como este. Assim, cobre todo o programa da disciplina com bastante detalhe.

4.2 - Bibliografia complementar

- [Ram00] *José Carlos Ramalho, Anotação Estrutural de Documentos e sua Semântica*, Edição, Escola de Engenharia - Universidade do Minho, Julho de 2000.
ISBN :
Tese de doutoramento
- [HM2001] *Elliott Rusty Harold, W. Scott Means, XML in a Nutshell*, First Edition Edição, O'Reilly, 2001.
ISBN :
- [Tid2001] *Doug Tidwell, XSLT*, Edição, O'Reilly, Agosto de 2001.
ISBN :

5 - MATERIAL PEDAGÓGICO

5.1 - Referências Web:

- [Site XML da O'Reilly](#)
- [Recursos do JCR de apoio à disciplina](#)

6 - SUMÁRIOS

Semana 1

Aula nº 1

Tipo	T1
Data	2005-03-05
Hora	9h-11h
Docente	José Carlos Ramalho
Sumário	

Um pouco de história: SGML, HTML, XML, SML.

As origens destas normas. Como é que elas se inter-relacionam.

O conceito de documento estruturado. Os vários tipos de anotação e sua evolução nos últimos tempos. As vantagens e as desvantagens da Anotação Descritiva. Como é composta uma linguagem de anotação e qual a sua aplicação. O conceito de meta-linguagem de anotação: o SGML e o XML.

As componentes dum documento XML: a declaração, o DTD e a instância. Documento válido versus documento bem-estruturado. Elementos estruturais dum documento XML: elementos, atributos e entidades. Alguns exemplos de documentos XML: o soneto, os sumários, algumas obras literárias e informação importada de bases de dados.

Semana 2

Aula nº 2

Tipo	T1
Data	2005-03-12
Hora	9h-11h
Docente	José Carlos Ramalho
Sumário	

Modelo de Dados para documentos XML: a árvore documental abstracta.

Seleccção e navegação na árvore documental abstracta: o XPath como linguagem de query.

Semana 3

Aula nº 3

Tipo	T1
Data	2005-03-19
Hora	9h-11h
Docente	José Carlos Ramalho
Sumário	

Transformação de documentos como transformação de árvores abstractas. O XSL e a transformação declarativa de documentos XML. Estrutura duma especificação em XSL e respectivo modelo de processamento. Os vários operadores e funções do XSL.

Semana 4

Aula nº 4

Tipo	T1
Data	2005-04-02
Hora	9h-11h
Docente	José Carlos Ramalho
Sumário	

Introdução às transformações com várias travessias sobre a árvore documental abstracta.

Resolução de alguns pequenos exemplos.

Semana 5

Aula nº 5

Tipo	T1
-------------	----

Data 2005-04-09
Hora 9h-11h
Docente José Carlos Ramalho
Sumário

Introdução aos Schemas: DTDs versus Schemas. Comparação de Schemas com gramáticas independentes de contexto. Desenvolvimento de Schemas para alguns casos de estudo.

Semana 6

Aula nº 6

Tipo T1
Data 2005-04-16
Hora 9h-11h
Docente José Carlos Ramalho
Sumário

Xslides: desenvolvimento de aplicação XML para suportar apresentações multimédia. Desenvolvimento integrado do XML Schema, instâncias exemplo e algumas transformações.

Semana 7

Aula nº 7

Tipo T1
Data 2005-04-30
Hora 9h-11h
Docente José Carlos Ramalho
Sumário

XPhotoAlbum: desenvolvimento de uma aplicação XML para suportar a criação e manutenção de albuns fotográficos.

Introdução ao processamento em simultâneo de várias árvores documentais abstractas com geração de várias saídas em simultâneo.

Semana 8

Aula nº 8

Tipo	T1
Data	2005-05-07
Hora	9h-11h
Docente	José Carlos Ramalho
Sumário	

Mapas Virtuais: desenvolvimento de uma aplicação XML para suportar a especificação de grafos e respectiva geração de websites para posterior navegação. Utilização de eixos de navegação em XPath para gerar as ligações entre as várias páginas HTML geradas.

Semana 9

Aula nº 9

Tipo	T1
Data	2005-05-13
Hora	18h-20h
Docente	José Carlos Ramalho
Sumário	

Finalização das aplicações iniciadas nas últimas aulas: Xslides, XPhotoAlbum e Mapas Virtuais.

Semana 10

Aula nº 10

Tipo	T1
Data	2005-05-14
Hora	9h-11h
Docente	José Carlos Ramalho
Sumário	

Finalização das aplicações iniciadas nas últimas aulas: Xslides, XPhotoAlbum e Mapas Virtuais.

Semana 11

Aula nº 11

Tipo	T1
Data	2005-05-21
Hora	9h-11h
Docente	José Carlos Ramalho
Sumário	

Introdução ao XSLFO: o que são *flow-objects*, instalação do processador, construção de pequenos exemplos.

Conceitos gerais: geometria de página, regiões, blocos, construção de páginas.

Semana 12

Aula nº 12

Tipo	T1
Data	2005-05-28
Hora	9h-11h
Docente	José Carlos Ramalho
Sumário	

Introdução a mais componentes do XSLFO.

Desenvolvimento da transformação para PDF para a aplicação *Relatório*.
