

# Processamento Estruturado de Documentos 2001

## LESI + LMCC (5º ano)

Ficha Teórico-Prática nº3

Data: 7 de Novembro de 2001

**Exercício 1 (Arqueossítios)** *Vá à página de Material Didático e descarregue o documento XML que descreve os arqueossítios do Noroeste Português.*

*Realize então, cada uma das seguintes tarefas:*

a) *Crie uma stylesheet XSL para transformar estes documentos em HTML:*

- *lembre-se da aplicação dum stylesheet a um documento XML:*

```
sabcmd arq.xsl arq.xml
```

b) *Vá buscar à página da cadeira a script atrib-id.pl e aplique-a ao documento XML.*

c) *Altere a stylesheet desenvolvida na primeira alínea de modo a criar um índice remissivo para cada um dos arqueossítios no topo da página HTML. Use os atributos acrescentados na alínea anterior para os links.*

**Exercício 2 (Arqueossítios e XPath)** *Para, nesta aula, poderes testar os teus conhecimentos de XPath abstraíndo-te da restante complexidade dum stylesheet XSL, foi desenvolvida uma metodologia: foi criado um pequeno DTD para documentos que irão conter expressões XPath e que depois de processados em conjunto com uma determinada instância documental dão como resultado um novo documento XML composto pelas subárvores resultantes da aplicação das expressões XPath à instância documental.*

*O DTD para documentos XPath é simples:*

```
<!ELEMENT query-set (query+)>  
<!ELEMENT query (subquery?)>  
<!ATTLIST query xpexp CDATA #REQUIRED>
```

*O atributo xpexp do elemento query deverá ser a expressão XPath a ser calculada (neste atributo só poderá o subconjunto do XPath suportado no atributo match das templates, se a query exigir mais complexidade coloque a parte restante no sub-elemento subquery).*

*Para se obter o resultado do cálculo dum expressão XPath sobre uma instância documental (por exemplo um poema) é necessário seguir os dois passos seguintes:*

1. *usar um gerador de stylesheets XSL que aplicado ao documento com as expressões XPath vai gerar uma stylesheet XSL que calcula essas expressões; foi desenvolvido um gerador, que não é mais nem menos do que uma stylesheet XSL — o xpath2.xsl; disponível na página da disciplina; para se obter a stylesheet em causa é necessário executar a linha seguinte:*

```
sabcmd xpath2.xsl q1.xml > q1.xsl
```

2. *a seguir utiliza-se a stylesheet gerada para processar a instância e obter o resultado do cálculo das expressões:*

```
sabcmd q1.xsl documento.xml > q1-res.xml
```

Tente resolver os seguintes exercícios respectivos ainda ao documento dos arqueossítios:

- a) Crie um documento baseado no DTD acima com expressões XPath que respondam às seguintes questões:
- os arqueossítios pertencentes ao concelho de Amares
  - os arqueossítios que tenham ligações a termos do dicionário
  - a identificação dos arqueossítios que se encontram a uma latitude superior a 180
  - a identificação dos arqueossítios que tenham o nome "Martins" na bibliografia
- b) Faça o mesmo com o documento dos sumários de PED2000 (também estão na página) e obtenha uma resposta para as seguintes questões:
- todas as aulas teóricas
  - todas as aulas práticas
  - a lista das datas das aulas
  - o primeiro parágrafo do sumário de cada aula
- c) Crie uma script XML::DT que transforme os documentos criados em páginas HTML que têm no início um índice de queries e que permitem navegar nos resultados. Analise previamente os documentos resultado deste sistema e pense no mapeamento mais correcto dos elementos deste tipo de documento para elementos HTML.

**Exercício 3 (Sumários)** Especifique uma stylesheet XSL que transforma o documento dos sumários numa página HTML atendendo aos seguintes requisitos: 1) As aulas teóricas aparecem primeiro ordenadas cronologicamente; 2) As aulas práticas aparecem a seguir também ordenadas cronologicamente; 3) no topo da página aparece um índice cronológico de aulas.

**Exercício 4 (Bibliografia)** Descarregue o ficheiro relativo à Bibliografia do JCR da página da disciplina e tente resolver as alíneas seguintes:

- a) Desenvolva uma stylesheet XSL, e uma script XML::DT para gerar uma versão HTML do documento atendendo aos seguintes pormenores:
- Cada entrada deverá estar individualizada das restantes (linhas horizontais, espaçamento, ...).
  - No topo da página, deverá aparecer um índice (disposto horizontalmente ou verticalmente) que permita um acesso directo a cada entrada. Utilize o identificador da entrada.
  - Tente criar um índice de autores em que a cada autor está associada a lista de identificadores das respectivas entradas. Pense na script ou na stylesheet, faça a que lhe parecer mais fácil. Depois tente fazer a outra.
- b) Se ainda tiver tempo especifique um DTD para este tipo de documentos.

**Exercício 5 (Submissão de TPs)** Foi montado um sistema para submissão electrónica de trabalhos práticos. Este sistema guarda a informação enviada num ficheiro XML com a seguinte estrutura:

```
...
<grupo>
  <aluno1>
    <numero1>19670</numero1>
```

```
<nome1>Alexandra Pinto</nome1>
<curso1>LESI</curso1>
</aluno1>
<aluno2>
  <numero2>22639</numero2>
  <nome2>André Sousa</nome2>
  <curso2>LESI</curso2>
</aluno2>
<aluno3>
  <numero3></numero3>
  <nome3></nome3>
  <curso3></curso3>
</aluno3>
<titulo>Processador genérico de documentos estruturados (XML)</titulo>
<file>19670_22639_.pli</file>
</grupo>
...
```

*Agora é necessário criar uma stylesheet que processe esta informação e crie uma página HTML onde seja fácil consultar a informação das várias submissões:*

- *Crie uma tabela.*
- *Cada grupo será uma linha da tabela.*
- *A informação dos elementos dum grupo poderá ficar numa sub-tabela.*
- *Crie um link para o documento enviado.*