
Integração de Dados

2021/22

XPath

Para testar expressões Xpath

- Usar um editor XML: editix
- Ferramentas on-line
 - <http://codebeautify.org/Xpath-Tester>
 - <http://www.freeformatter.com/xpath-tester.html>

DOM = Document Object Model

- **Fornecer a representação estrutural de documentos HTML e XML**
- **Define a forma como a estrutura pode ser acessada por programas e scripts**
- **Possibilita a modificação na estrutura do documento, do estilo e do conteúdo**

XML DOM

- **Em DOM, tudo num documento XML é visto como um **nodo**:**
 - o documento XML é um **nodo documento**
 - um elemento XML é um **nodo elemento**
 - o texto dos elementos XML são **nodos de texto**
 - um atributo é um **nodo atributo**
 - um comentário é um **nodo comentário**.
- **As linguagens de programação fornecem API para DOM**
 - funções que permitem percorrer a árvore (documento)
 - Em Java: o API é o JDOM

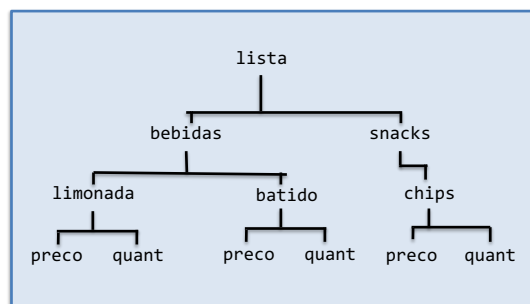
XML DOM

- Representação em árvore
- Relações entre nodos da árvore
 - parent (pai)
 - child (filho)
 - sibling (irmão)
 - ancestor(ascendente)
 - descendant (descendente)

XML DOM – exemplo

```
<lista>
  <bebidas>
    <limonada>
      <preco>2.50</preco>
      <quant>20</quant>
    </limonada>
    <batido>
      <preco>1.50</preco>
      <quant>10</quant>
    </batido>
  </bebidas>
  <snacks>
    <chips>
      <preco>4.50</preco>
      <quant>60</quant>
    </chips>
  </snacks>
</lista>
```

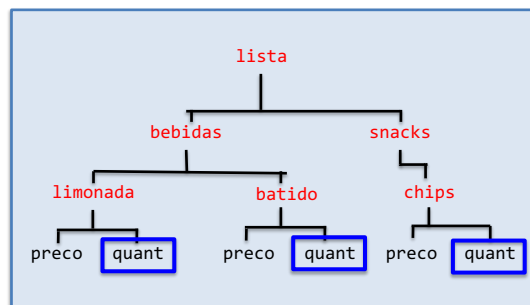
ÁRVORE



XML DOM – exemplo

```
<lista>
  <bebidas>
    <limonada>
      <preco>2.50</preco>
      <quant>20</quant>
    </limonada>
    <batido>
      <preco>1.50</preco>
      <quant>10</quant>
    </batido>
  </bebidas>
  <snacks>
    <chips>
      <preco>4.50</preco>
      <quant>60</quant>
    </chips>
  </snacks>
</lista>
```

ANCESTORS de <quant>? ANTEPASSADOS



@Anabela Simões

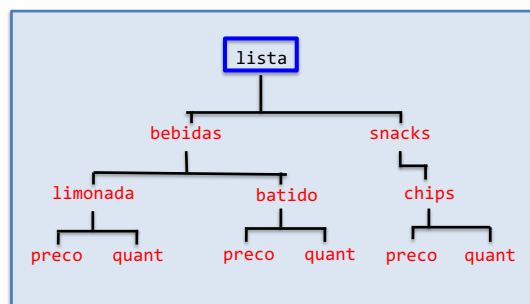
Integração de Dados – 2021/22

7

XML DOM – exemplo

```
<lista>
  <bebidas>
    <limonada>
      <preco>2.50</preco>
      <quant>20</quant>
    </limonada>
    <batido>
      <preco>1.50</preco>
      <quant>10</quant>
    </batido>
  </bebidas>
  <snacks>
    <chips>
      <preco>4.50</preco>
      <quant>60</quant>
    </chips>
  </snacks>
</lista>
```

DESCENDANTS de <lista>? DESCENDENTES



@Anabela Simões

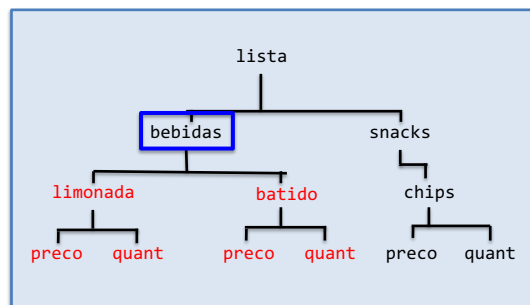
Integração de Dados – 2021/22

8

XML DOM – exemplo

```
<lista>
  <bebidas>
    <limonada>
      <preco>2.50</preco>
      <quant>20</quant>
    </limonada>
    <batido>
      <preco>1.50</preco>
      <quant>10</quant>
    </batido>
  </bebidas>
  <snacks>
    <chips>
      <preco>4.50</preco>
      <quant>60</quant>
    </chips>
  </snacks>
</lista>
```

DESCENDANTS de <bebidas>? DESCENDENTES



@Anabela Simões

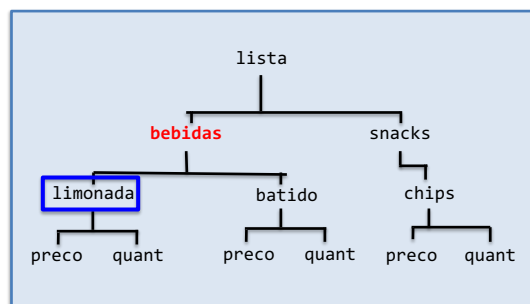
Integração de Dados – 2021/22

9

XML DOM – exemplo

```
<lista>
  <bebidas>
    <limonada>
      <preco>2.50</preco>
      <quant>20</quant>
    </limonada>
    <batido>
      <preco>1.50</preco>
      <quant>10</quant>
    </batido>
  </bebidas>
  <snacks>
    <chips>
      <preco>4.50</preco>
      <quant>60</quant>
    </chips>
  </snacks>
</lista>
```

PARENT de <limonada>? PAI



@Anabela Simões

Integração de Dados – 2021/22

10

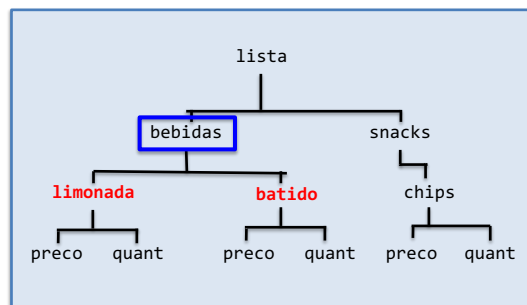
XML DOM – exemplo

```

<lista>
  <bebidas>
    <limonada>
      <preco>2.50</preco>
      <quant>20</quant>
    </limonada>
    <batido>
      <preco>1.50</preco>
      <quant>10</quant>
    </batido>
  </bebidas>
  <snacks>
    <chips>
      <preco>4.50</preco>
      <quant>60</quant>
    </chips>
  </snacks>
</lista>

```

**CHILD de <bebidas>?
FILHO**



@Anabela Simões

Integração de Dados – 2021/22

11

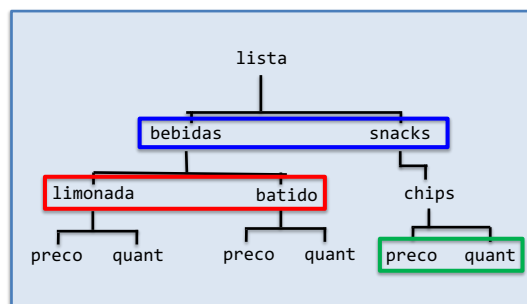
XML DOM – exemplo

```

<lista>
  <bebidas>
    <limonada>
      <preco>2.50</preco>
      <quant>20</quant>
    </limonada>
    <batido>
      <preco>1.50</preco>
      <quant>10</quant>
    </batido>
  </bebidas>
  <snacks>
    <chips>
      <preco>4.50</preco>
      <quant>60</quant>
    </chips>
  </snacks>
</lista>

```

**SIBLING: mesmo nível e mesmo *parent*
IRMÃO:**



@Anabela Simões

Integração de Dados – 2021/22

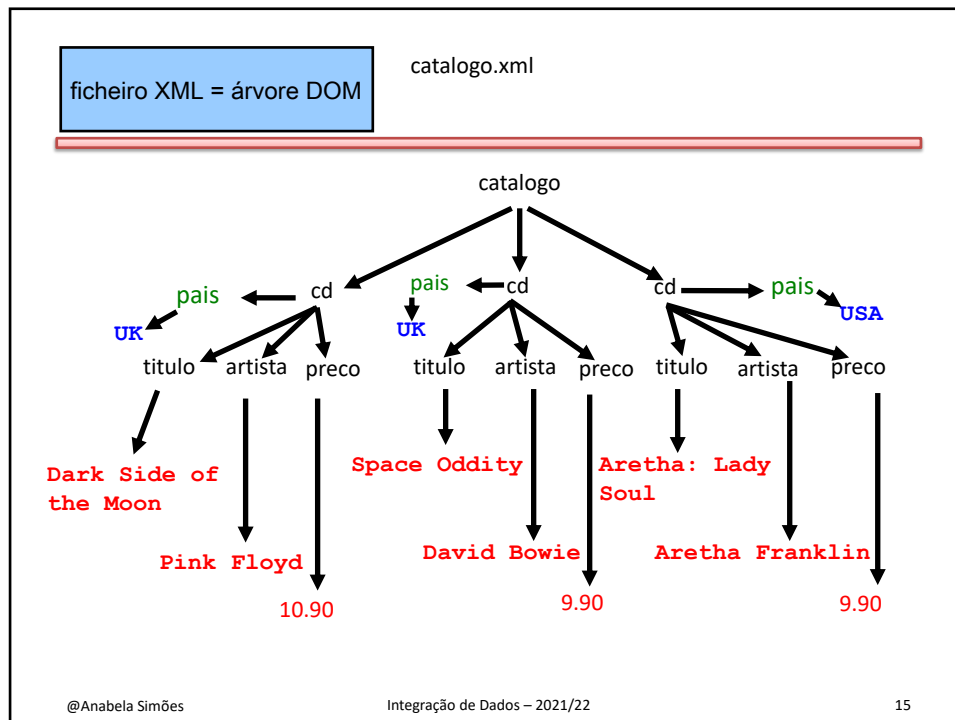
12

Extraír Dados de um ficheiro XML

- XML armazena dados logo é necessário extraí-los para usar em diversas aplicações
- Dados do XML podem ser extraídos:
 - por um programa
 - usando a linguagem XPath
- XPath é utilizada noutras linguagens:
 - XML Schema (XSD)
 - XSLT
 - XQuery
 - ...

catalogo.xml

```
<?xml version="1.0" encoding="UTF-8"?>
<catalogo>
  <cd pais="UK">
    <titulo>Dark Side of the Moon</titulo>
    <artista>Pink Floyd</artista>
    <preco>10.90</preco>
  </cd>
  <cd pais="UK">
    <titulo>Space Oddity</titulo>
    <artista>David Bowie</artista>
    <preco>9.90</preco>
  </cd>
  <cd pais="USA">
    <titulo>Aretha: Lady Soul</titulo>
    <artista>Aretha Franklin</artista>
    <preco>9.90</preco>
  </cd>
</catalogo>
```



XPath: introdução

- É uma linguagem que constrói expressões de pesquisa baseadas em caminhos
 - um ficheiro XML **F** é uma **árvore**
 - uma **expressão E** especifica os caminhos possíveis em **F**
 - a expressão **E** devolve os **nodos** ou **valores** de **F** que são encontrados pelos(s) caminho(s) especificado(s)
- **As expressões XPath especificam:**
 - navegação em ficheiros (caminhos)
 - testes (filtros) aos nodos

XPath: introdução

Uma expressão XPath pode devolver:

- uma lista de nodos (elementos, atributos)

ou

- uma lista de valores (numérico, string, ...)

XPath: introdução

Pesquisa: Todos os CDs do ficheiro

Devolve uma lista de elementos <cd>

```
Item 1: <cd pais="UK">
      <titulo>Dark Side of the Moon</titulo>
      <artista>Pink Floyd</artista>
      <preco>10.90</preco>
    </cd>
Item 2: <cd pais="UK">
      <titulo>Space Oddity</titulo>
      <artista>David Bowie</artista>
      <preco>9.90</preco>
    </cd>
Item 3: <cd pais="USA">
      <titulo>Aretha: Lady Soul</titulo>
      <artista>Aretha Franklin</artista>
      <preco>9.90</preco>
    </cd>
```

XPath: introdução

Pesquisa: Todos os artistas do ficheiro

Elementos?

```
Item 1: <artista>Pink Floyd</artista>
Item 2: <artista>David Bowie</artista>
Item 3: <artista>Aretha Franklin</artista>
```

Valores?

```
Item 1: Pink Floyd
Item 2: David Bowie
Item 3: Aretha Franklin
```

Expressões XPath

- Todas as expressões XPath seleccionam nodos a partir do início de pesquisa indicado:
 - **Absolutas**: um **/** no início da expressão indica que a pesquisa inicia na raiz

/ step/step/...
 - **Relativas**: a pesquisa inicia a partir do nodo actual

step/step/...

Expressões XPath

- /** no início de uma expressão XPath representa a raiz do documento
- /** entre nomes de elementos representa uma relação parent/child (pai/filho)
- //** representa uma relação ancestor-descendant (antepassado-descendente), i.e, procura em todo o documento, o elemento que se segue a //
- .** representa **o nodo actual**
- ..** representa **o pai do nodo actual**
- *** representa qualquer elemento
- @** assinala um atributo
- |** procura em caminhos alternativos

@Anabela Simões

Integração de Dados – 2021/22

21

Expressões Xpath – alguns exemplos

/catalogo/cd/titulo

/catalogo//artista

//preco

//cd/@pais

@Anabela Simões

Integração de Dados – 2021/22

22

Expressões XPath: condições ou predicados

- **elemento[condição]**
 - especifica uma condição
 - a condição é uma subexpressão Xpath cujo caminho deve ser contruido a partir de **elemento**
- **elemento[N]**
 - indica a posição de elemento.
 - Devolve elemento nessa posição

@Anabela Simões

Integração de Dados – 2021/22

23

Expressões XPath: alguns exemplos

```
/catalogo/cd[@pais="UK"]/preco
```

```
//cd[preco < 10]
```

```
//cd[preco < 10]/artista
```

```
//cd[2] | /cd[3]
```

```
/catalogo/cd[2]/artista
```

@Anabela Simões

Integração de Dados – 2021/22

24

Expressões XPath: operadores

- para criar condições é necessário usar operadores de comparação:

=	igual
!=	diferente
> < <= >=	outros, comparação
not(...)	negação
and	conjunção
or	disjunção

@Anabela Simões

Integração de Dados – 2021/22

25

Expressões XPath: funções

<http://saxon.sourceforge.net/saxon7.9.1/functions.html>

- **count(lista)**
 - Conta nº de elementos devolvidos por uma expressão Xpath
 - Recebe uma expressão Xpath
 - Devolve um número inteiro
- **contains(s1,s2)**
 - Verifica se a string s2 está contida no string s1
 - Recebe duas Strings (podem resultar de expressões XPath)
 - Devolve true ou false
- **text()**
 - obtém o texto de um elemento folha (NÃO SERVE PARA ATRIBUTOS)
 - Coloca-se no final de uma expressão XPath que termine num element folha
- **data(elem)**
 - obtém o texto de um elemento ou atributo
 - Recebe uma expressão XPath que devolva um valor atómico
 - Devolve o conteúdo da expressão XPath sem o XML

@Anabela Simões

Integração de Dados – 2021/22

26

Expressões XPath: funções – alguns exemplos

```
count(/catalogo/cd)
```

```
count(//artista)
```

```
/catalogo/cd[count(titulo)=1]/artista
```

```
//cd[contains(titulo, "Moon")]/artista
```

```
//cd/titulo[contains(., "Moon")]/../artista
```

```
/catalogo/cd/titulo/text()
```

```
data(//artista)
```

@Anabela Simões

Integração de Dados – 2021/22

27

Expressões XPath: funções

<http://saxon.sourceforge.net/saxon7.9.1/functions.html>

- **last()**
 - Devolve o número de items do nodo actual
 - Colocado como predicado de uma expressão Xpath
- **position()**
 - devolve posição do nodo actual
 - Colocado como predicado de uma expressão Xpath
- **sum(n1, n2, ...)**
 - Devolve a soma dos argumentos
 - Recebe uma lista numérica – pode ser resultado de uma expressão XPath
- **min(n1, n2, ...)**
 - Devolve o valor mais baixo de uma lista numérica
 - Recebe uma lista numérica – pode ser resultado de uma expressão XPath
- **max(n1, n2, ...)**
 - Devolve o valor mais alto de uma lista numérica
 - Recebe uma lista numérica – pode ser resultado de uma expressão XPath

@Anabela Simões

Integração de Dados – 2021/22

28

Expressões XPath: funções – alguns exemplos

`/catalogo/cd[last()]`

`/catalogo/cd[position()<=2]/titulo`

`sum(//preco)`

`min(//ano)`

`max(//preco)`

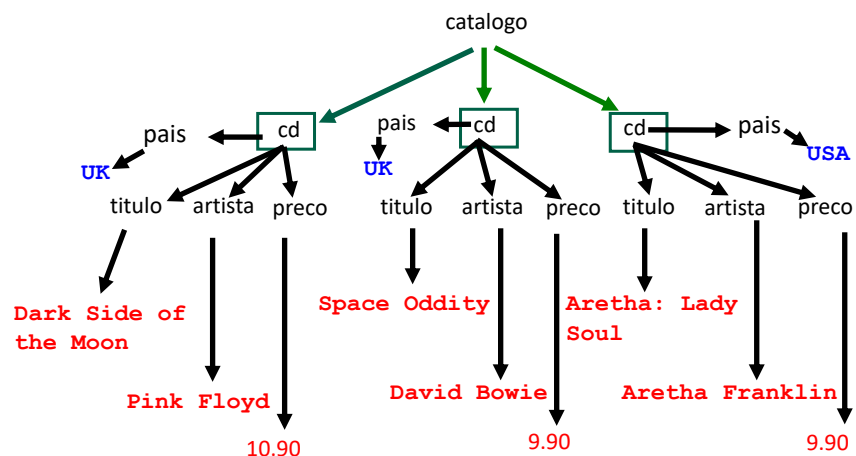
@Anabela Simões

Integração de Dados – 2021/22

29

`/catalogo/cd`

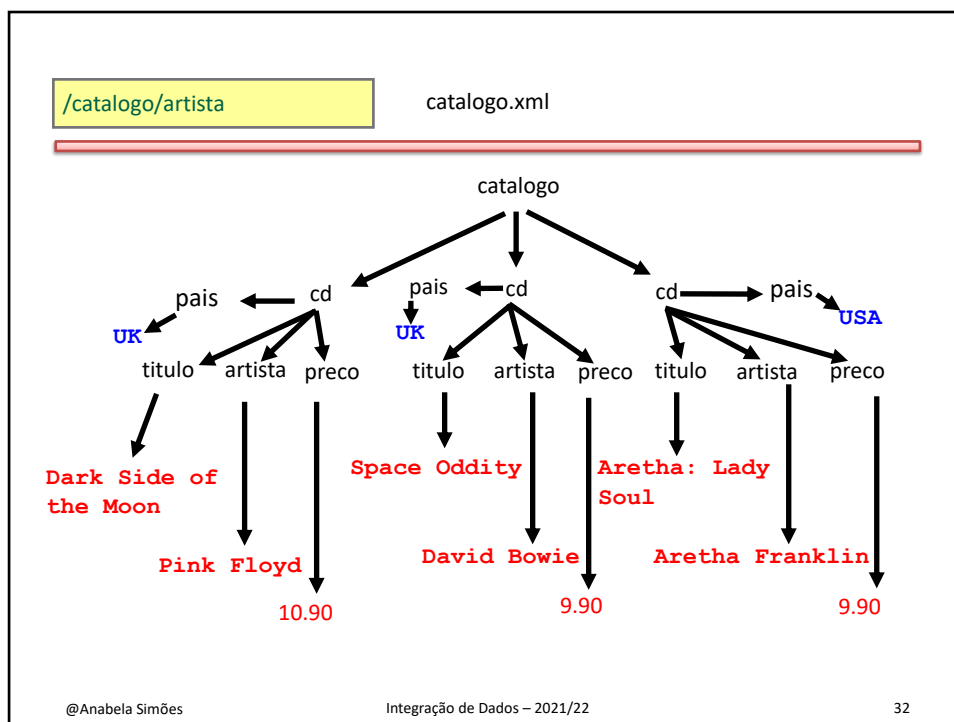
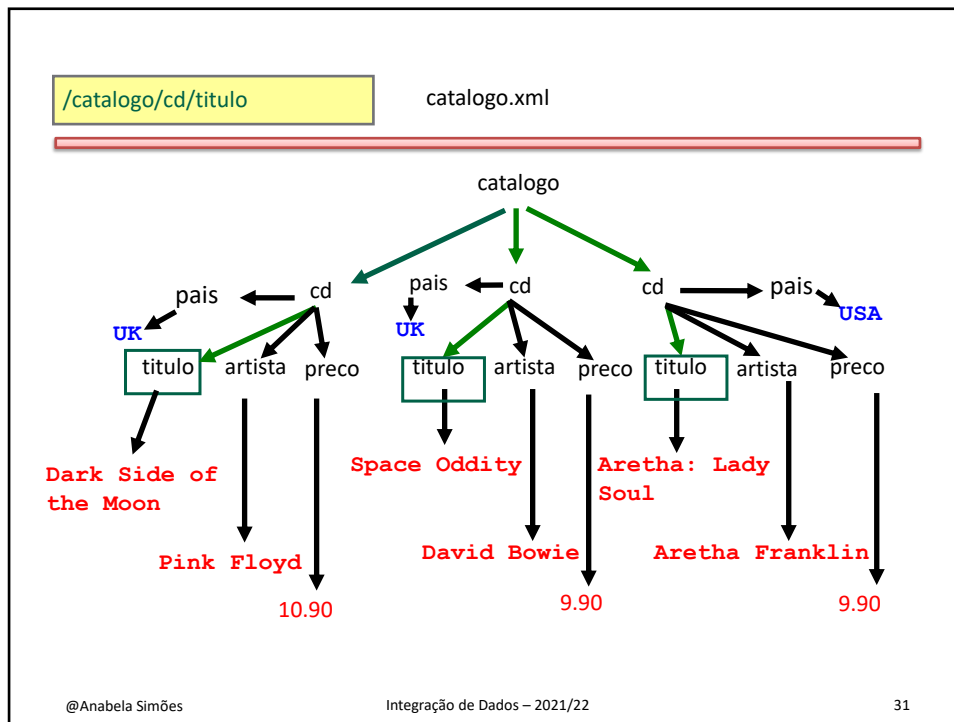
catalogo.xml

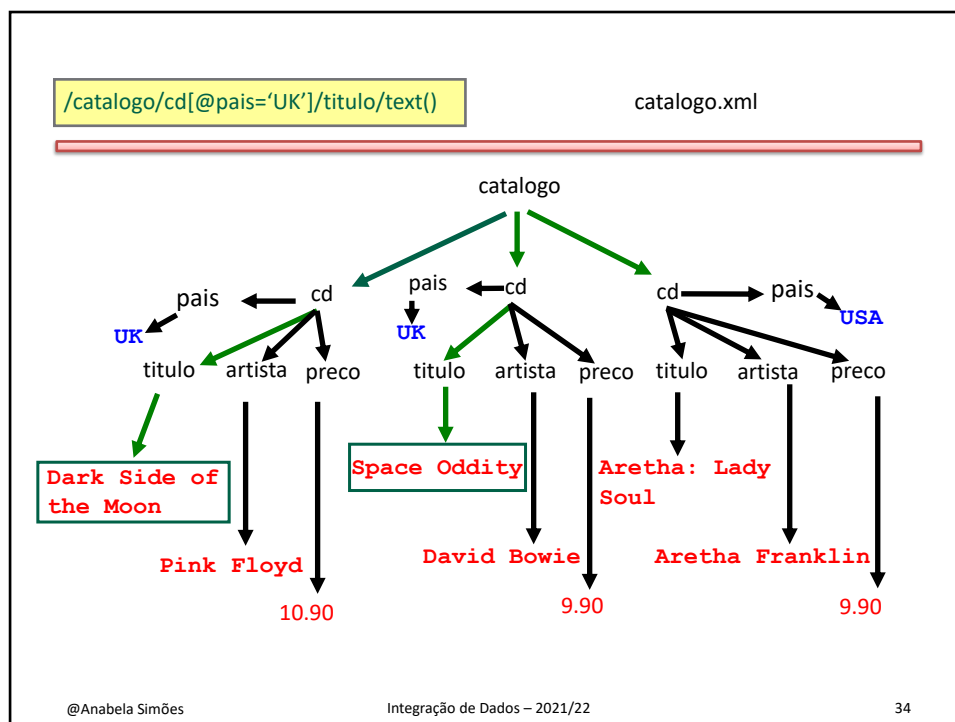
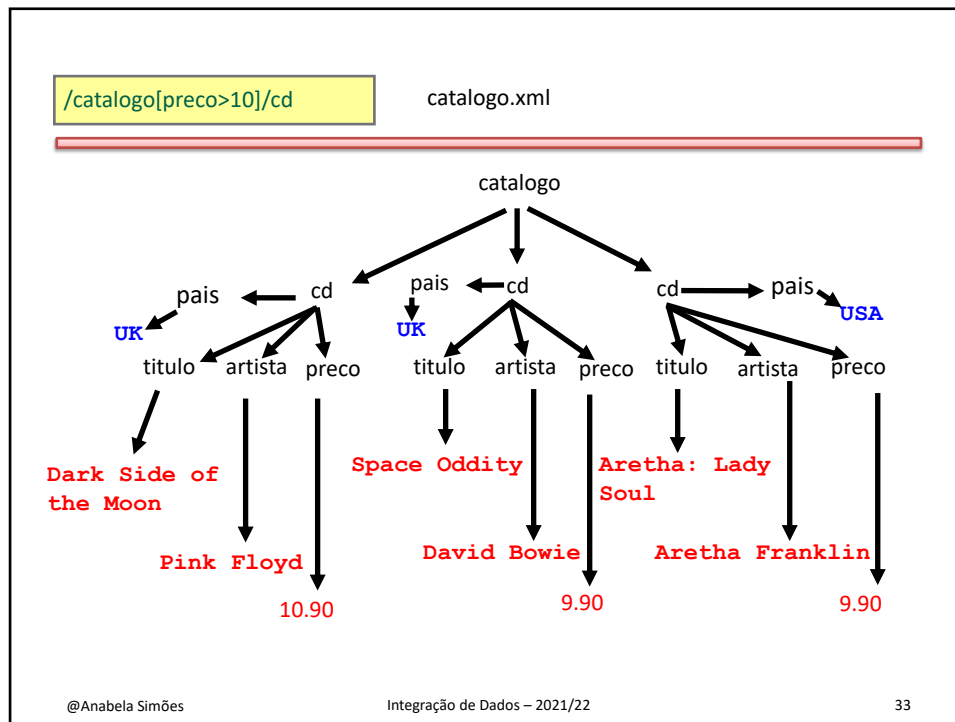


@Anabela Simões

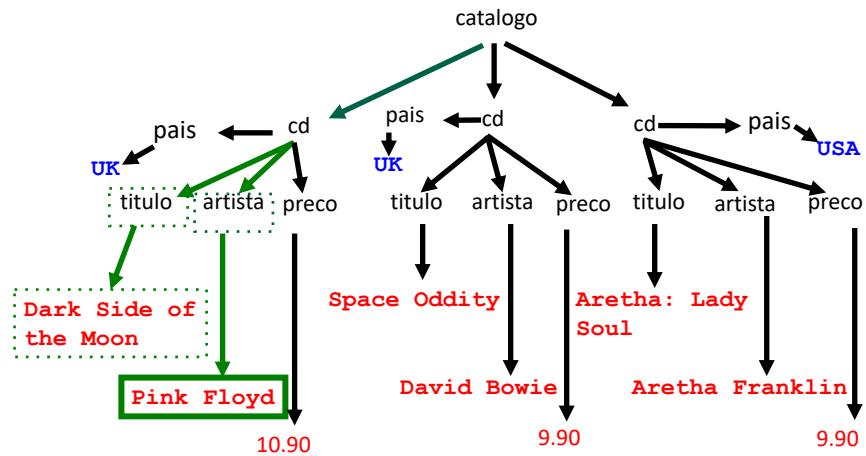
Integração de Dados – 2021/22

30





```
/catalogo/cd[@pais="UK"]/titulo[contains(., "Moon")]/../artista/text()
```



@Anabela Simões

Integração de Dados – 2021/22

35

Expressões XPath: exemplos

Indique o resultado de cada uma das expressões XPath

```

/catalogo/cd/preco
/catalogo/cd/preco/text()

/catalogo/cd[preco<10]
/catalogo/cd/preco[.<10]

/catalogo//titulo/
//titulo

/catalogo/cd[1]
/catalogo/cd[last()]

sum(//preco)
count(//cd[preco<10])

```

```

max(//preco)

//cd/preco[.<10]/../titulo/text()
//cd[preco<10]/titulo/text()
//cd[preco=max(//preco)]/titulo/text()

//titulo | //preco

/catalogo/cd[@pais="UK"]
data(/catalogo/cd/@pais)

/catalogo/cd/*
//*[preco=9.90]
//*[preco=9.90]/*

```

@Anabela Simões

Integração de Dados – 2021/22

36

Expressões XPath: exemplos

Escreva uma expressão XPath para cada uma das seguintes pesquisas

- Encontrar o(s) preço(s) do(s) cd(s) de David Bowie?
- Encontrar o(s) país(es) do(s) cd(s) do artista Aretha Franklin
- Encontrar o(s) país(es) do(s) cd mais caro(s)
- Encontrar o(s) artista(s) do(s) cd(s) mais barato(s)
- Encontrar nodos descendentes do(s) cd(s) dos USA

@Anabela Simões

Integração de Dados – 2021/22

37

Ficheiro XML: **imobiliaria.xml**

<pre> <imobiliaria > <imoveis> <imovel tipo="Apt" tipologia="T2"> <localizacao>Ilhavo, Aveiro</localizacao> <area uni="m2">100</area> <preço>75000</preço> <extras> <extra>Ar condicionado</extra> <extra>Aquecimento central</extra> <extra>Garagem</extra> </extras> </imovel> <imovel tipo="Apt" tipologia="T3"> <localizacao>Aveiro</localizacao> <area uni="m2">140</area> <preço>12000</preço> <extras> <extra>Garagem</extra> </extras> </imovel> <imovel tipo="Moradia" tipologia="T5"> <localizacao>Coimbra</localizacao> <area uni="m2">300</area> <preço>475000</preço> <extras> <extra>Ar condicionado</extra> <extra>Aquecimento central</extra> <extra>Piscina</extra> <extra>Garagem</extra> </extras> </imovel> ... </imoveis> </imobiliaria > </pre>	<pre> <vendedores> <persoa classe="vip"> <nome>Ana Martins</nome> <contacto tipo="tlm">966644567</contacto> <localidade>Coimbra</localidade> <idade>22</idade> </persoa> <persoa classe="nivel 1"> <nome>Joana Martins</nome> <contacto tipo="tlm">911199911</contacto> <contacto tipo="fixo">244555000</contacto> <localidade>Leiria</localidade> <idade>44</idade> </persoa> <persoa classe="nivel 2"> <nome>Jorge Melo</nome> <contacto tipo="tlm">965559993</contacto> <contacto tipo="tlm">928883339</contacto> <localidade>Lisboa</localidade> <idade>32</idade> </persoa> <persoa classe="nivel 2"> <nome>Isabel Martins</nome> <contacto tipo="tlm">965559993</contacto> <localidade>Porto</localidade> <idade>27</idade> </persoa> </vendedores> </imobiliaria> </pre>
---	---

Expressões XPath: Exercícios

Analise o ficheiro XML anterior e escreva uma expressão XPath para cada uma das seguintes pesquisas

- Nomes (elemento nome) dos vendedores do tipo “vip”
- Localidade(s) (texto) onde vivem os vendedores de apelido **Martins**
- Qual a média das idades dos vendedores que vivem em **Coimbra**?
- Primeiro vendedor (elemento) registado no ficheiro com dois contactos telefónicos
- Atributo **tipologia** dos imóveis com mais de 100 m2 de área e situados em qualquer zona de **Aveiro**
- Qual a tipologia do apartamento com a menor área?
- Número de imóveis com mais do que dois extras

Expressões XPath

- **A navegação usando a sintaxe básica do XPath permite apenas navegar no sentido:**

- **Descendente**

Usando os caminhos formados com / ou //

- **Ascendente**

Usando os caminhos formados com / ou // e o operador ..

- **Lateral: é necessário usar os eixos de navegação**

XPath – eixos de navegação

Expressões XPath

- Sintaxe XPath completa

nome_eixo::nome_nodo[predicado]

nome_eixo: define um conjunto de nodos relativamente ao nodo atual

nome_nodo: nodo a selecionar com o eixo indicado

predicado: filtro aplicado ao *nome_nodo*

por exemplo:

Sem eixos: **//cd[preco > 10]/artista**

Com eixos:

descendant::cd[child::preco > 10]/child::artista

XPath: Eixos de navegação

Eixos:

- devolvem um conjunto de nodos relativos ao nodo atual
- alguns eixos são opcionais
- indicam o sentido da navegação a partir do nodo indicado
- **Até agora usámos os eixos de forma omitida/abreviada**
 - **child::** (default, se nada for indicado)
 - **descendant::** (//)
 - **attribute::** (@)
 - **self::** (.)
 - **parent::** (..)

@Anabela Simões

Integração de Dados – 2021/22

43

Expressões XPath: mais eixos

Nome do eixo	Resultado
ancestor	accede aos antepassados do nodo atual
ancestor-or-self	accede aos antepassados do nodo atual e o nodo atual
descendant	accede aos descendentes do nodo atual
descendant-or-self	accede aos descendentes do nodo atual e o nodo atual
following	accede a todos os nodos que existam no documento depois da tag de fecho do nodo atual
following-sibling	accede aos irmãos depois do nodo atual
preceding	accede a todos os nodos que existam no documento antes da tag de abertura do nodo atual
preceding-sibling	accede aos irmãos antes do nodo atual

@Anabela Simões

Integração de Dados – 2021/22

44

Expressões XPath equivalentes (catalogo.xml)

```
/child::catalogo/descendant::cd
```

```
/catalogo//cd
```

```
/catalogo/descendant-or-self::*
```

```
/catalogo/* | /catalogo
```

```
/catalogo/child::cd[attribute::pais="UK"]/preco
```

```
/catalogo/cd[@pais="UK"]/preco
```

```
/catalogo/cd[attribute::pais="USA"]/preco/parent::cd/artista
```

```
/catalogo/cd[attribute::pais="USA"]/preco/./artista
```

@Anabela Simões

Integração de Dados – 2021/22

45

```
<lista>
  <bebidas>
    <limonada type="fresh">
      <preco>2.50</preco>
      <quant>20</quant>
    </limonada>
    <batido type="summer">
      <preco>1.50</preco>
      <quant>10</quant>
    </batido>
  </bebidas>
  <snacks>
    <chips>
      <preco>4.50</preco>
      <quant>60</quant>
    </chips>
    <amendoins>
      <preco>3.50</preco>
      <quant>100</quant>
    </amendoins>
  </snacks>
</lista>
```

limonada.xml

@Anabela Simões

Integração de Dados – 2021/22

46

Expressões XPath equivalentes (limonada.xml)

```
/lista/bebidas/limonada/preco/../../batido
/lista/bebidas/limonada/preco/parent::limonada/parent::bebidas/batido
/lista/bebidas/limonada/preco/ancestor::bebidas/batido
```

Não omitir nenhum eixo?

```
/child::lista/child::bebidas/child::limonada/child::preco/
parent::limonada/parent::bebidas/child::batido
```

```
/lista/bebidas/limonada/preco/../../quant
```

```
/lista/bebidas/limonada/preco/parent::limonada/quant
```

Não omitir nenhum eixo?

```
/child::lista/child::bebidas/child::limonada/child::preco/parent::
:limonada/child::quant
```

@Anabela Simões

Integração de Dados – 2021/22

47

Expressões XPath equivalentes (limonada.xml)

```
//bebidas/*[@type="fresh"]
```

```
descendant::bebidas/*[@type="fresh"]
```

Não omitir nenhum eixo?

```
descendant::bebidas/child::*[attribute::type="fresh"]
```

```
//bebidas/*[quant > 10 and @type="fresh"]/preco
```

R:

Não omitir nenhum eixo?

R: ...

@Anabela Simões

Integração de Dados – 2021/22

48

receitas.xml

```

<receitas>
  <receita>
    <titulo>Panquecas</titulo>
    <ingrediente nome="farinha" quant="250" unid="g"/>
    <ingrediente nome="leite" quant="3" unid="dl"/>
    <ingrediente nome="margarina" quant="70" unid="g"/>
    <ingrediente nome="acucar" quant="10" unid="g"/>
    <ingrediente nome="ovos" quant="2"/>
    <preparacao>
      <passo>Numa tigela, bater os ovos, o açúcar, a margarina, o leite e a farinha.</passo>
      <passo>Bater bem até obter uma massa homogênea.</passo>
      <passo>Aquecer frigideira</passo>
      <passo>Deitar uma concha do preparado e deixar cozinhar.</passo>
      <passo>Virar com a ajuda de uma espátula e deixar alourar.</passo>
    </preparacao>
    <infos>Servir com doce, mel ou xarope.</infos>
    <nutricao calorias="612" gord="49"/>
  </receita>

  <receita>
    <titulo>Ovo cozido</titulo>
    <ingrediente nome="ovo" quant="1"/>
    <ingrediente nome="sal" quant="5" unid="g"/>
    <preparacao>
      <passo>Ferver água numa caçarola.</passo>
      <passo>Colocar o ovo na água.</passo>
      <passo>Deixar cozer durante 6 minutos.</passo>
      <passo>Retirar o ovo e servir descascado.</passo>
    </preparacao>
    <nutricao calorias="60" gord="15"/>
  </receita>
  ...
</receitas>

```

@Anabela Simões
Integração de Dados – 2021/22
49

Exercícios: o que devolvem as expressões?

descendant::ingrediente[@quant=5 or @quant=10]

Todos os elementos <ingrediente> que tenham o atributo quant igual a 5 ou 10

descendant::ingrediente[position()=2]

O 2º ingrediente que aparece no ficheiro

descendant::*[self::ingrediente][2]

O 2º ingrediente que aparece no ficheiro

/receitas/descendant::receita/child::ingrediente[position()=2]

O 2º ingrediente de cada receita

/receitas/descendant::receita[last()]/preceding::receita[2]

A antepenúltima receita

Exercícios: usar eixos de navegação

1. Todos os ingredientes excepto o primeiro da receita “Panquecas”
2. Preparação, nutrição e infos da receita “Panquecas”
3. Lista de passos exceptuando o último da receita “Paquecas”
4. Lista de todos os ingredientes que se seguem ao leite de todas as receitas que usem leite

@Anabela Simões

Integração de Dados – 2021/22

51

Exercícios: construa expressões **usando eixos** que encontrem

O penúltimo ingrediente da receita das **Panquecas**

`/receitas/receita[titulo='Panquecas']/ingrediente[last()]/preceding::ingrediente[1]`

O antepenúltimo passo da receita das **Panquecas**

`/receitas/receita[titulo='Panquecas']//passo[last()]/preceding::passo[2]`

Titulo de todas as receitas, excepto a última

`/receitas/receita[last()]/preceding-sibling::receita/titulo`

Quantas receitas levam mais do que dois ingredientes?

`count(/receitas/child::receita[count(child::ingrediente)>2])`

@Anabela Simões

Integração de Dados – 2021/22

52

Exercícios: usando eixos, construa expressões XPath que encontrem

Título das receitas que incluam o açúcar na lista de ingredientes

`descendant::receita/ingrediente[attribute::nome='acucar']/../titulo`

Título das receitas que tenham não tenham açúcar na lista de ingredientes

`descendant::receita[count(ingrediente[attribute::nome="acucar"])=0]/titulo`

O segundo atributo do segundo ingrediente de todas as receitas

`/child::receitas/child::receita/child::ingrediente[2]/attribute::*[2]`

Qual o título (texto) das receitas com menos de 3 passos de preparação?

`/descendant::receita[count(preparacao/child::passo) <3]/child::titulo/text()`

@Anabela Simões

Integração de Dados – 2021/22

53

```
<?xml version="1.0" encoding="UTF-8"?>
<catalogo>
  <livro categoria="Culinária">
    <titulo lang="pt">Doçaria conventual</titulo>
    <autor>Maria Silva</autor>
    <ano>2005</ano>
    <preco>30.00</preco>
  </livro>
  <livro categoria="Juvenil">
    <titulo lang="pt">Harry Potter 1</titulo>
    <autor>J. K. Rowling</autor>
    <ano>2005</ano>
    <preco>29.99</preco>
  </livro>
  <livro categoria="Juvenil">
    <titulo lang="pt">Harry Potter 2</titulo>
    <autor>J. K. Rowling</autor>
    <ano>2005</ano>
    <preco>19.99</preco>
  </livro>
  <livro categoria="WEB">
    <titulo lang="en">XML Bible</titulo>
    <autor>James McGovern</autor>
    <autor>Per Bothner</autor>
    <autor>Kurt Cagle</autor>
    <autor>James Linn</autor>
    <autor>Vaidyanathan Nagarajan</autor>
    <ano>2003</ano>
    <preco>49.99</preco>
  </livro>
  <livro categoria="WEB">
    <titulo lang="en">Learning XML</titulo>
    <autor>Erik T. Ray</autor>
    <ano>2001</ano>
    <preco>39.95</preco>
  </livro>
</catalogo>
```

catalogo_livros.xml

@Anabela Simões

Integração de Dados – 2021/22

54

XPath (Exemplos 1/5)

- Selecionar todos os elementos títulos:

```
//titulo
/catalogo/titulo
descendant::titulo
```

- Selecionar o título (conteúdo) do 1º livro:

```
/catalogo/livro[1]/titulo/text()
/child::catalogo/child::livro[1]/child::titulo/text()
```

- Qual o preço do penúltimo livro:

```
/catalogo/livro[last()-1]/preco
/catalogo/livro[last()]/preceding-sibling::livro[1]/preco
/child::catalogo/child::livro[last()]/preceding-
sibling::livro[1]/child::preco
```

@Anabela Simões

Integração de Dados – 2021/22

55

XPath (Exemplos 2/5)

- Preço do livro mais caro publicado em 2005
- O preço do livro mais caro escrito por J. K. Rowling
- Extrair, em letras maiúsculas, o título do livro mais antigo do catálogo
- Média dos preços dos livros escritos por J. K. Rowling
- Categoria do primeiro livro que aparece antes da categoria Web

@Anabela Simões

Integração de Dados – 2021/22

56

XPath (Exemplos 3/5) – o que devolve?

```
//livro[count(autor)=5]/autor[contains(., 'Linn')]/../@categoria
```

- escreva expressão equivalente sem abreviatura dos eixos:

@Anabela Simões

Integração de Dados – 2021/22

57

Ficheiro XML: **imobiliaria.xml**

<pre><imobiliaria > <imoveis> <imovel tipo="Apt" tipologia="T2"> <localizacao>Ilhavo, Aveiro</localizacao> <area uni="m2">100</area> <preço>75000</preço> <extras> <extra>Ar condicionado</extra> <extra>Aquecimento central</extra> <extra>Garagem</extra> </extras> </imovel> <imovel tipo="Apt" tipologia="T3"> <localizacao>Aveiro</localizacao> <area uni="m2">140</area> <preço>12000</preço> <extras> <extra>Garagem</extra> </extras> </imovel> <imovel tipo="Moradia" tipologia="T5"> <localizacao>Coimbra</localizacao> <area uni="m2">300</area> <preço>475000</preço> <extras> <extra>Ar condicionado</extra> <extra>Aquecimento central</extra> <extra>Piscina</extra> <extra>Garagem</extra> </extras> </imovel> </imoveis> ... </imobiliaria ></pre>	<pre><vendedores> <persoa classe="vip"> <nome>Ana Martins</nome> <contacto tipo="tlm">966644567</contacto> <localidade>Coimbra</localidade> <idade>22</idade> </persoa> <persoa classe="nivel 1"> <nome>Joana Martins</nome> <contacto tipo="tlm">911199911</contacto> <contacto tipo="fixo">244555000</contacto> <localidade>Leiria</localidade> <idade>44</idade> </persoa> <persoa classe="nivel 2"> <nome>Jorge Melo</nome> <contacto tipo="tlm">965559993</contacto> <contacto tipo="tlm">928883339</contacto> <localidade>Lisboa</localidade> <idade>32</idade> </persoa> <persoa classe="nivel 2"> <nome>Isabel Martins</nome> <contacto tipo="tlm">965559993</contacto> <localidade>Porto</localidade> <idade>27</idade> </persoa> </vendedores> </imobiliaria></pre>
---	---

Expressões XPath: Exercícios

Analise o ficheiro XML anterior e escreva uma expressão XPath **usando eixos de navegação** para cada uma das seguintes pesquisas

- Área e localização de todos os imóveis abaixo de 100000
- Os 3 imóveis anteriores ao último
- Localidades, contactos e nomes dos vendedores com idade abaixo de 30 anos