

Lista de Exercícios

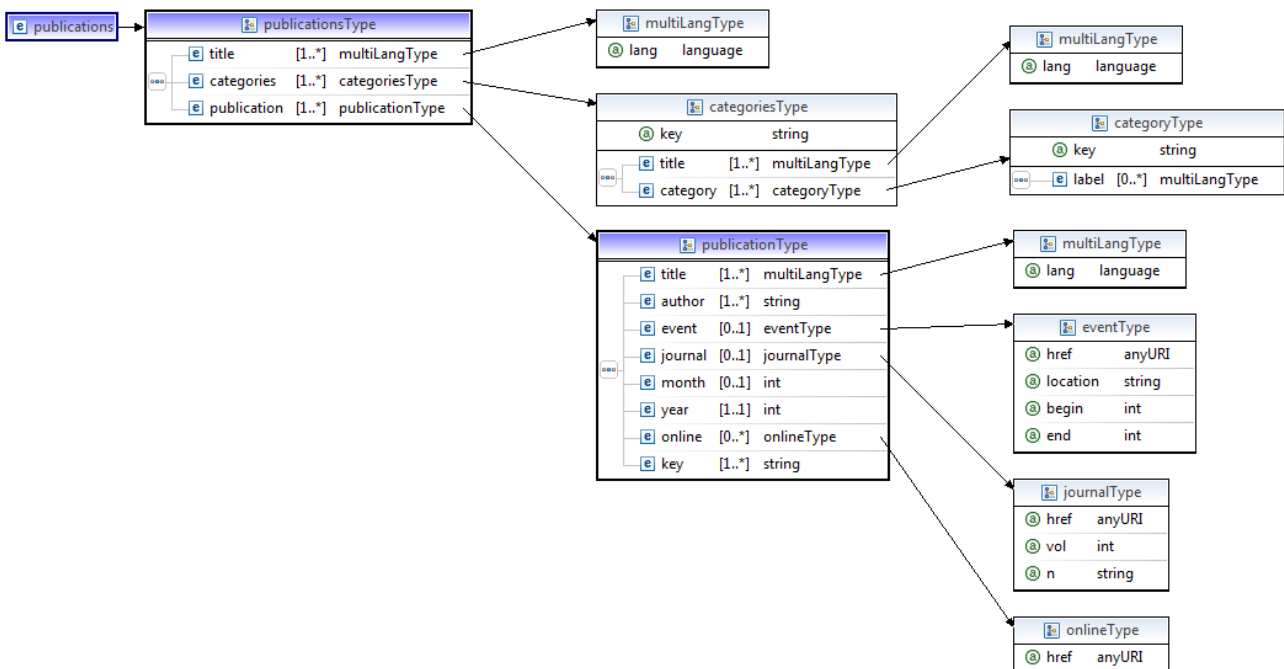
MC536 - Bancos de Dados: Teoria e Prática
 Instituto de Computação
 Universidade Estadual de Campinas

Banco de Dados na Web
 XML, XPath, XQuery
 2013
 André Santanchè

Considere a representação gráfica a seguir de um esquema XML referente a publicação de artigos em conferências e revistas científicas (*journals*). O elemento `<publications>` define um título `<title>` geral, um conjunto de categorias `<categories>` (em que as publicações estão categorizadas) e as publicações - elementos `<publication>`. Diversos elementos, como o `<title>`, são definidos como `multiLangType`, que indica que além da `string` de conteúdo, têm um atributo `@lang` que define em que língua o conteúdo está.

O conjunto de categorias define a chave que será usada para classificar as categorias `@catkey` e um título descritivo do conjunto de categorias `<title>`. Cada categoria `<category>` dentro do conjunto define uma chave de identificação `@key` e rótulo. Há um tipo de categoria especial `year` que será usada para agrupar publicações por ano.

Cada publicação `<publication>` define título `<title>`, um conjunto de autores `<author>`, mês `<month>` e ano `<year>` em que foi publicada e chaves `@key`, que associam a publicação a categorias. Uma publicação terá um elemento `<event>` se foi publicada em um evento (e.g., conferência ou workshop); para cada evento é definido o link `@href` para a sua página, onde ele aconteceu (`@location`) e qual os dias de início `@begin` e fim `@end`. Se, por outro lado, a publicação for em uma revista científica, um elemento `<journal>` indicará o link `@href` da revista, o volume `@vol` e número `@n` em que o artigo foi publicado. Quando as publicações estão disponíveis online, um ou mais links para os respectivos arquivos são providos pelo elemento `<online>` através do atributo `@href`.



O esquema e arquivo XML com dados estão disponíveis online em:

- <http://www.ic.unicamp.br/~santanche/publications/publications.xsd>
- <http://www.ic.unicamp.br/~santanche/publications/publications.xml>

Questão 1 - XPath

Elabore expressões Xpath que representam as questões a seguir (você pode testá-las em <http://www.online-toolz.com/tools/xpath-editor.php>):

- a) Elementos completos das categorias classificadas por “subject” (definido pelo @catkey).

```
//categories[@catkey="subject"]/category  
ou
```

```
/publications/categories[@catkey="subject"]/category
```

- b) Texto dos rótulos em português das categorias classificadas por “subject”.

```
//categories[@catkey="subject"]/category/label[@lang="pt-BR"]/text()
```

- c) Texto dos títulos das publicações posteriores ao ano de 2010.

```
//publication[year>2010]/title/text()  
ou
```

```
/publications/publication[year>2010]/title/text()
```

- d) Publicações de André Santanchè.

```
//publication[author="André Santanchè"]
```

- e) Último autor de cada publicação em que aparece André Santanchè.

```
//publication[author="André Santanchè"]/author[last()]
```

Questão 2 - XQuery

Elabore expressões XQuery para as consultas especificadas a seguir (você pode testá-las em <http://www.zorba-xquery.com/html/demo>):

Para as respostas a seguir considere sempre:

```
let $publicadoc := doc('http://www.ic.unicamp.br/~santanche/publications/publications.xml')
```

- a) Retorne quantas publicações são posteriores ao ano de 2011.

```
return count($publicadoc//publication[year>2010])
```

- b) Para cada publicação, retorne um XML dentro com elemento raiz <summary>, contendo um elemento <title> com o título da publicação e um elemento <nauthors> com o número de autores.

```
for $p in ($publicadoc//publication)  
return <summary>  
    <title>{data($p/title)}</title>  
    <nauthors>{count($p/author)}</nauthors>  
</summary>
```

- c) Se o título da publicação não tiver atributo @lang, haverá um único título, caso contrário, pode haver títulos em várias línguas, mas sempre haverá um título em inglês “en-US”. Refine a consulta anterior para que retorne o único título (se não houver @lang) ou o título em inglês se houver mais de uma língua.

```
for $p in ($publicadoc//publication)  
return <summary>  
    <title>{  
        if ($p/title/@lang)
```

```

        then data($p/title[@lang='en-US'])
        else data($p/title)
    }</title>
    <nauthors>{count($p/author)}</nauthors>
</summary>

```

d) Retorne a categoria cujo <label> em inglês seja 'e-Science Domain'.

```

for $c in ($publicadoc//categories/category)
where $c/label[@lang='en-US'] = 'e-Science Domain'
return $c

```

e) Retorne as publicações associadas à categoria cujo <label> em inglês seja 'e-Science Domain'. A associação entre o label e a key da categoria deve ser feita na consulta.

```

for $c in ($publicadoc//categories/category),
    $p in ($publicadoc//publication)
where $c/label[@lang='en-US'] = 'e-Science Domain' and
    $p/key = $c/@key
return $p

```

f) Para cada publicação, retorne um XML dentro com elemento raiz <publ>, contendo um elemento <title> com o título da publicação e um elemento <category> com o label em inglês da publicação.

```

for $c in ($publicadoc//categories/category),
    $p in ($publicadoc//publication)
where $p/key = $c/@key
return <publ>
    <title>{data($p/title)}</title>
    <category>{data($c/label[@lang='en-US'])}</category>
</publ>

```

g) Retorne o título das publicações agrupados por categoria da seguinte maneira:

- um elemento <group> agrupa cada categoria e dentro dele haverá:
 - um elemento <category> com o label em inglês da categoria
 - um conjunto de elementos <ptitle> com os títulos das publicações daquela categoria.

```

for $c in ($publicadoc//categories[@catkey = 'subject']/category)
return <group>
    <category>{data($c/label[@lang='en-US'])}</category>
    {for $p in ($publicadoc//publication)
     where $p/key = $c/@key
     return <ptitle>{data($p/title)}</ptitle>}
</group>

```