



Instituto Superior de Engenharia de Coimbra

**Curso Técnico Superior e Profissional –
Tecnologias e Programação de Sistemas de Informação**

Leandro Adão Fidalgo | a21270093

Pedro dos Santos Alves | a21270246

**Integração de Dados
Trabalho Prático**

Coimbra, 20 de janeiro de 2019

Índice

1. Introdução	3
2. Definir o Esquema Global	4
3. Câmaras	4
4. Contratos	5
5. Gerar Ficheiros XML	5
6. Definir Pesquisas	12
7. Gerar Ficheiros Output	14
8. Conclusão	15

1. Introdução

O presente relatório descreve o projeto desenvolvido pelos alunos: Leandro Fidalgo e Pedro Alves, no âmbito da disciplina de Integração de Dados, no Curso Técnico Superior Profissional de Tecnologias e Programação de Sistemas de Informação do Instituto Superior de Engenharia de Coimbra. O propósito deste trabalho é consolidar todos os conhecimentos adquiridos no decorrer do semestre.

O objetivo do trabalho consiste na criação de uma aplicação integradora que apresente uma vista unificada de informação relativa a Câmaras Municipais Portuguesas da Região Centro, contendo a informação detalhada e relativa aos contratos das Câmaras Municipais Portuguesas da Região Centro. As fontes de dados a usar são: <http://www.base.gov.pt/Base/pt/ResultadosPesquisa?type=contratos&query=adjudicanteid%3D> e <https://www.anmp.pt/anmp/pro/mun1/mun101w2.php?dis=06>. Com estas duas fontes de dados o objetivo consiste em efetuar a integração de dados e construir dois ficheiros XML que agreguem a informação de forma organizada e coerente.

Para isso, é necessário realizar as seguintes tarefas: definir o esquema global (análise do ficheiro `camaras.txt`), implementar os *wrappers* (realizar expressões regulares para obter as informações necessárias), gerar os ficheiros XML em função do esquema global, definir pesquisas sobre os ficheiros XML criados nos pontos anteriores e sobre as fontes de dados originais, editar e eliminar informações e por fim gerar ficheiros de output.

Durante o processo de realização deste projeto foram surgindo algumas dúvidas, que conseguiram ser superadas através do auxílio do professor da disciplina e das fichas de trabalho por ele disponibilizadas.

2. Definir o Esquema Global

Na fase inicial do projeto, é necessário criar os URL's para aceder às câmaras e aos contratos através da web. Constatou-se que a parte inicial do url é sempre contante nos dois sites: <https://www.anmp.pt/anmp/pro/mun1/mun101w2.php?dis=06> e <http://www.base.gov.pt/Base/pt/ResultadosPesquisa?type=contratos&query=adjudicanteid%3D>. E foi nos dado pelo professor os seguintes nif's das câmaras que tinha de retirar informações.

3. Câmaras

Deve ser obtido através do website <https://www.anmp.pt/anmp/pro/mun1/mun101w2.php?dis=06> a seguinte informação: Nome do Município, Nome do Presidente da Câmara, E-mail, Site Institucional, Telefone, Número de Freguesias, Área, Número Habitantes, Brasão, Feriado e Nome do Presidente da Assembleia.

Expressão regular para o nome do município:

```
class="sel3">([a-zA-Z\sóãá-])[\s]*</div>
```

Expressão regular para o nome do presidente da câmara:

```
class="f3">([a-zA-Zó-íãéÉú\ç\á\Ã\Í+), Presidente da C
```

Expressão regular para o e-mail:

```
<a href="mailto:([a-zA-Z@\-\.]+)
```

Expressão regular para o site institucional:

```
gina\sWeb:\s([a-zA-Z:\-\.]+)
```

Expressão regular para o telefone:

```
Telefone:\s([0-9\s]+)
```

Expressão regular para o número de freguesias:

```
(\d+) freguesias
```

Expressão regular para a área:

```
(\d+,\d+) km2
```

Expressão regular para o número de habitantes:

Código geração do XML (*camaras.xml*):

Página 6 de 15

Página 7 de 15

Ficheiro **camaras.xml**:

```
<?xml version="1.0" encoding="utf-8"?>
<camaras>
  <camara id="id0">
    <municipio>Arganil</municipio>
    <presidenteC>Luis Paulo Carreira Fonseca</presidenteC>
    <presidenteA>Ricardo João Barata Pereira Alves</presidenteA>
    <email>geral@cm-arganil.pt</email>
    <site>http://www.cm-arganil.pt</site>
    <telefone>235 200 150</telefone>
    <numfreguesias>14</numfreguesias>
    <area>333,0</area>
    <numhabitantes>12 145</numhabitantes>
    <feriado>2018-09-07</feriado>
    <brasao>https://www.anmp.pt/image/munap/M3300b.jpg</brasao>
  </camara>
  <camara id="id1">
    <municipio>Cantanhede</municipio>
    <presidenteC>Maria Helena Rosa de Teodósio</presidenteC>
    <presidenteA>João Pais Moura</presidenteA>
    <email>geral@cm-cantanhede.pt</email>
    <site>http://www.cantanhedeonline.pt</site>
    <telefone>231 410 100</telefone>
    <numfreguesias>14</numfreguesias>
    <area>391,0</area>
    <numhabitantes>36 595</numhabitantes>
    <feriado>2018-07-25</feriado>
    <brasao>https://www.anmp.pt/image/munap/M3060b.jpg</brasao>
  </camara>
  <camara id="id2">
    <municipio>Coimbra</municipio>
    <presidenteC>Manuel Machado</presidenteC>
    <presidenteA>Luís Marinho</presidenteA>
    <email>geral@cm-coimbra.pt</email>
    <site>http://www.cm-coimbra.pt</site>
    <telefone>239 857 500</telefone>
    <numfreguesias>18</numfreguesias>
    <area>319,0</area>
    <numhabitantes>143 396</numhabitantes>
    <feriado>2018-07-04</feriado>
    <brasao>https://www.anmp.pt/image/munap/M3000b.jpg</brasao>
  </camara>
</camaras>
```


Ficheiro **camaras.xsd**:

```
<?xml version="1.0" encoding="UTF-8"?>

<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">

  <xsd:element name="municipio" type="xsd:string" />

  <xsd:element name="presidenteC" type="xsd:string" />

  <xsd:element name="presidenteA" type="xsd:string" />

  <xsd:element name="feriado" type="xsd:string" />

  <xsd:element name="email" type="xsd:string" />

  <xsd:element name="site" type="xsd:string" />

  <xsd:element name="telefone">
    <xsd:simpleType>
      <xsd:restriction base="xsd:string">
        <xsd:pattern value="23[0-9]\s*[0-9]{3}\s*[0-9]{3}" />
      </xsd:restriction>
    </xsd:simpleType>
  </xsd:element>

  <xsd:element name="numfreguesias" type="xsd:string" />

  <xsd:element name="area" type="xsd:string" />

  <xsd:element name="numhabitantes" type="xsd:string" />

  <xsd:element name="brasao">
    <xsd:simpleType>
      <xsd:restriction base="xsd:string">
        <xsd:pattern value="https://\www\anmp\pt/image/munap/M[0-9]+b.jpg" />
      </xsd:restriction>
    </xsd:simpleType>
  </xsd:element>

  <xsd:attribute name="id" type="xsd:ID" />

  <xsd:element name="camaras">
    <xsd:complexType>
      <xsd:sequence>

        <xsd:element name="camara" minOccurs="1" maxOccurs="unbounded">
          <xsd:complexType>
            <xsd:sequence>

              <xsd:element ref="municipio" minOccurs="1" maxOccurs="1" />

              <xsd:element ref="presidenteC" minOccurs="1" maxOccurs="1" />

              <xsd:element ref="presidenteA" minOccurs="0" maxOccurs="1" />

              <xsd:element ref="email" minOccurs="1" maxOccurs="1" />

              <xsd:element ref="site" minOccurs="1" maxOccurs="1" />

              <xsd:element ref="telefone" minOccurs="1" maxOccurs="1" />

              <xsd:element ref="numfreguesias" minOccurs="1" maxOccurs="1" />

              <xsd:element ref="area" minOccurs="1" maxOccurs="1" />

              <xsd:element ref="numhabitantes" minOccurs="1" maxOccurs="1" />

              <xsd:element ref="feriado" minOccurs="1" maxOccurs="1" />

              <xsd:element ref="brasao" minOccurs="1" maxOccurs="1" />

            </xsd:sequence>
            <xsd:attribute ref="id" use="required" />
          </xsd:complexType>
        </xsd:element>

      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>

</xsd:schema>
```

[illegible]

Ficheiro contratos.xml:

```
<?xml version="1.0" encoding="utf-8"?>
<contratos>
  <municipio nome="Arganil">
    <nif>506833232</nif>
    <contrato id="id0">
      <objcontrato>INTERVENÇÕES VÁRIAS EM INFRAESTRUTURAS HIDRÁULICAS NA EN342</objcontrato>
      <preco>122.932,85 €</preco>
      <datapub>15-01-2019</datapub>
      <adjudicatario>CIPIRIANO PEREIRA DE CARVALHO & FILHOS LDA</adjudicatario>
    </contrato>
    <contrato id="id1">
      <objcontrato>Prestação de serviço para a execução de produção de cartografia...</objcontrato>
      <preco>14.580,00 €</preco>
      <datapub>10-01-2019</datapub>
      <adjudicatario>GEOLAYER - GEOTECNIA E SERVIÇOS, LDA</adjudicatario>
    </contrato>
    <contrato id="id2">
      <objcontrato>Demolição de Imóveis em Benfeita, Vinhô, Pombeiro da Beira, Corgas,...</objcontrato>
      <preco>109.610,00 €</preco>
      <datapub>02-01-2019</datapub>
      <adjudicatario>Construções Alfredo Rodrigues José, Lda.</adjudicatario>
    </contrato>
    <contrato id="id3">
      <objcontrato>Ajuste direto para Fornecimento e Montagem de Equipamento para a...</objcontrato>
      <preco>14.780,20 €</preco>
      <datapub>28-12-2018</datapub>
      <adjudicatario>Elifri-Electrónica e Frio Industrial, Lda</adjudicatario>
    </contrato>
    <contrato id="id4">
      <objcontrato>Concurso Público Para Aquisição de Serviços de Ensaios Laboratoriais Acreditados...</objcontrato>
      <preco>66.080,00 €</preco>
      <datapub>28-12-2018</datapub>
      <adjudicatario>Itecons- Instituto de investigação e desenvolvimento tecnológico par a construção, energia, ambiente e sustentabilidade</adjudicatario>
    </contrato>
    <contrato id="id5">
      <objcontrato>Prestação de serviços para a instalação e substituição de contadores...</objcontrato>
      <preco>22.267,20 €</preco>
      <datapub>28-12-2018</datapub>
      <adjudicatario>CUNHA & ANTUNES, LDA</adjudicatario>
    </contrato>
  </municipio>
</contratos>
```

Ficheiro contratos.xsd:

```
<?xml version="1.0" encoding="utf-8"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">

  <xsd:element name="nif">
    <xsd:simpleType>
      <xsd:restriction base="xsd:string">
        <xsd:pattern value="[0-9]{9}" />
      </xsd:restriction>
    </xsd:simpleType>
  </xsd:element>

  <xsd:element name="objcontrato" type="xsd:string" />

  <xsd:element name="preco" type="xsd:string" />

  <xsd:element name="datapub">
    <xsd:simpleType>
      <xsd:restriction base="xsd:string">
        <xsd:pattern value="([0-2][0-9]{3}[01])-(0[0-9]|1[0-2])-(19|20)[0-9]{2}" />
      </xsd:restriction>
    </xsd:simpleType>
  </xsd:element>

  <xsd:element name="adjudicatario" type="xsd:string" />

  <xsd:attribute name="id" type="xsd:ID" />

  <xsd:attribute name="nome" type="xsd:string" />

  <xsd:element name="contratos">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element name="municipio" minOccurs="1" maxOccurs="unbounded">
          <xsd:complexType>
            <xsd:sequence>
              <xsd:element ref="nif" minOccurs="1" maxOccurs="1" />
            </xsd:sequence>
          </xsd:complexType>
        </xsd:element>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
</xsd:schema>
```

```
<xsd:element name="contrato" minOccurs="1" maxOccurs="unbounded">
  <xsd:complexType>
    <xsd:sequence>
      <xsd:element ref="objcontrato" minOccurs="1" maxOccurs="1" />
      <xsd:element ref="preco" minOccurs="1" maxOccurs="1" />
      <xsd:element ref="datapub" minOccurs="1" maxOccurs="1" />
      <xsd:element ref="adjudicatario" minOccurs="1" maxOccurs="1" />
    </xsd:sequence>
    <xsd:attribute ref="id" use="required" />
  </xsd:complexType>
</xsd:element>

</xsd:sequence>
<xsd:attribute ref="nome" use="required" />
</xsd:complexType>
</xsd:element>

</xsd:sequence>
</xsd:complexType>
</xsd:element>

</xsd:schema>
```

6. Definir Pesquisas

O programa deve permitir procurar contratos por data específica, procurar contratos por autor da publicação, procurar contratos por adjudicatário, procurar qual o contrato de maior valor de uma Câmara Municipal específica, procurar qual o contrato de maior valor de todas as Câmaras, introduzir uma câmara e obter todos os dados da mesma, o Top 5 das Câmara que gastaram mais com contratos, procurar Câmara por nome do presidente, efetuar pesquisas que combinem dois ou mais atributo, porém não conseguimos fazer o Top 5 das Câmara que gastaram mais com contratos, procurar qual o contrato de maior valor de uma Câmara Municipal específica e procurar qual o contrato de maior valor de todas as Câmaras.

Com o intuito de obter essas pesquisas foi necessário usar o *XPath* e a API JAXEN para fazer a ligação entre ambos. Para isso, foi necessário criar duas funções em Java(*pesquisaXPath()* e *listarResultado()*).

Função `listarResultado()`:

```
public static String listarResultado(List res) {
    StringBuilder lista = new StringBuilder();
    //lista = null;
    for (int i = 0; i < res.size(); i++) {
        if (res.get(i).getClass().getName().equals("org.jdom2.Element")) {
            Element x = (Element) res.get(i);
            // System.out.println("Elemento " + x.getName() + " = " + x.getValue());
            lista.append(x.getValue()).append("\n");
        }
        if (res.get(i).getClass().getName().equals("org.jdom2.Attribute")) {
            Attribute x = (Attribute) res.get(i);
            // System.out.println("Atributo " + x.getName() + " = " + x.getValue());
            lista.append(x.getValue() + "\n");
        }
        if (res.get(i).getClass().getName().equals("org.jdom2.Text")) {
            Text x = (Text) res.get(i);
            // System.out.println("Texto " + x.getValue());
            lista.append(x.getValue() + "\n");
        }
        if (res.get(i).getClass().getName().equals("java.lang.Double")) {
            Double x = (Double) res.get(i);
            System.out.println("Double = " + x);
            lista.append(x.toString() + "\n");
        }
    }
    if (res.isEmpty())
        return "Sem resultados";
    else

```

Função `pesquisaXPath()`:

```
public static List pesquisaXPath(Document doc, String expXPath) {
    try {
        XPath xp = XPath.newInstance(expXPath);

        //devolve todos os nós que respeitam a expressão XPath
        List resultado = xp.selectNodes(doc);
        return resultado;
    } catch (JDOMException ex) {
        Logger.getLogger(JaxenFunctions_XPATH.class.getName()).log(Level.SEVERE, null, ex);
    } catch (NoClassDefFoundError a) {
        System.out.println("Erro : Classe não definida" + a);
    }
    return null;
}
```

7. Gerar Ficheiros Output

O programa deve possibilitar ao utilizador gerar ficheiros de resultados. Estes ficheiros devem ser transformações dos ficheiros XML da vista global. Duas transformações obrigatórias: gerar um ficheiro HTML contendo o brasão de todos as Câmaras Municipais listadas no ficheiro XML e gerar um novo ficheiro XML que junte atributos dos dois ficheiros (contratos.xml e camaras.xml).

Ficheiro **brasões.xslt**:

```
<?xml version="1.0" encoding="utf-8"?>
<xsl:stylesheet version="1.0" xmlns:xsl="http://www.w3.org/1999/XSL/Transform">
  <xsl:output method="html" />

  <xsl:template match="camaras">
    <html>
      <head>
        <title>
          <xsl:text>Brasões</xsl:text>
        </title>
      </head>
      <body>
        <xsl:apply-templates select="camara" />
      </body>
    </html>
  </xsl:template>

  <xsl:template match="camara">
    <p>
      <xsl:text>Brasão de </xsl:text><xsl:value-of select="municipio" />
    </p>
    
    <hr />
  </xsl:template>
</xsl:stylesheet>
```

Função para juntar o ficheiro das câmaras e dos contratos:

```
private void FicheiroXMLActionPerformed(java.awt.event.ActionEvent evt) {GEN-FIRST:event_FicheiroXMLActionPerformed
// TODO add your handling code here:

// FICHEIRO XML
StringBuilder sb = new StringBuilder();
sb.append("<?xml version='1.0' encoding='utf-8'>").append('\n').append("<juntar>").append('\n');
try {
    BufferedReader BR = new BufferedReader(new FileReader(new File("fichs/camaras.xml")));
    String test;
    while ((test = BR.readLine()) != null) {
        if (!test.equals("<?xml version='1.0' encoding='utf-8'>")) {
            sb.append(test).append('\n');
        }
    }
    BR = new BufferedReader(new FileReader(new File("fichs/contratos.xml")));
    while ((test = BR.readLine()) != null) {
        if (!test.equals("<?xml version='1.0' encoding='utf-8'>")) {
            sb.append(test).append('\n');
        }
    }
    BR.close();
    sb.append("</juntar>");
    BufferedWriter BW = new BufferedWriter(new FileWriter(new File("fichs/transformado/juntar.xml")));
    BW.write(sb.toString());
    BW.close();
} catch (FileNotFoundException ex) {
    Logger.getLogger(NomeBemNotPodre.class.getName()).log(Level.SEVERE, null, ex);
} catch (IOException ex) {
    Logger.getLogger(NomeBemNotPodre.class.getName()).log(Level.SEVERE, null, ex);
}
}
//GEN-LAST:event_FicheiroXMLActionPerformed
```

8. Conclusão

A realização deste projeto permitiu consolidar diversas competências nomeadamente: a criação e a manipulação de ficheiros XML, exercitar as expressões regulares, realização de pesquisas de informação em ficheiros XML, usando XPath, efetuar validação de ficheiros XML usando DTD e/ou XSD, etc.

As principais dificuldades existentes no desenvolvimento do projeto foi a implementação de expressões regulares nos Contratos. Também surgiu dificuldade na interface gráfica do projeto aquando a apresentação da imagem. Estas dificuldades foram superadas com a ajuda do professor da disciplina.