Instituto Politécnico de Viana do Castelo

Tutorial 8
Java e XML para implementar o parsing de ficheiros XML – Em Ambiente Windows

20381 – João Sencadas - joaosencadas@ipvc.pt

20399 – Ruben Morim – rubenmorim@ipvc.pt

Orientador(es):

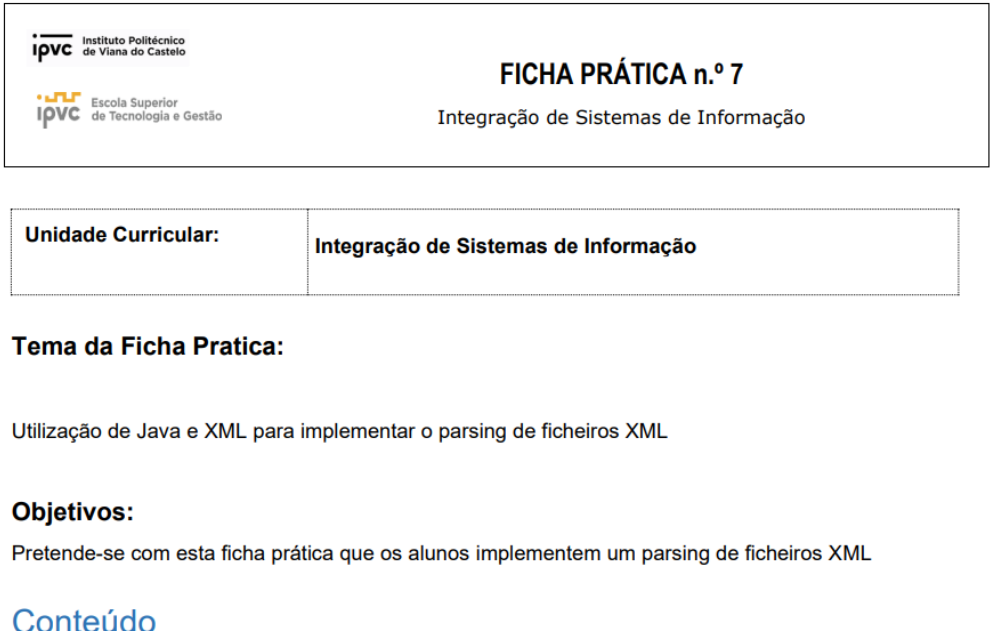
- Professor Doutor Jorge Ribeiro

Licenciatura em Engenharia Informática | Escola Superior de Tecnologia e Gestão| Unidade Curricular: **Integração de Sistemas** | Ano Letivo 2021/2022

Summary

1. Realização da ficha 7

Para este relatório iremos implementar o parsing de ficheiros XML, o objetivo é através de um servidor RPC enviar um ficheiro XML para o servidor o receba e faça o parsing desse ficheiro XML.



Instituto Politécnico de Viana do Castelo
Escola Superior de Tecnologia e Gestão

FICHA PRÁTICA n.º 7

Integração de Sistemas de Informação

Unidade Curricular:	**Integração de Sistemas de Informação**

Tema da Ficha Pratica:

Utilização de Java e XML para implementar o parsing de ficheiros XML

Objetivos:

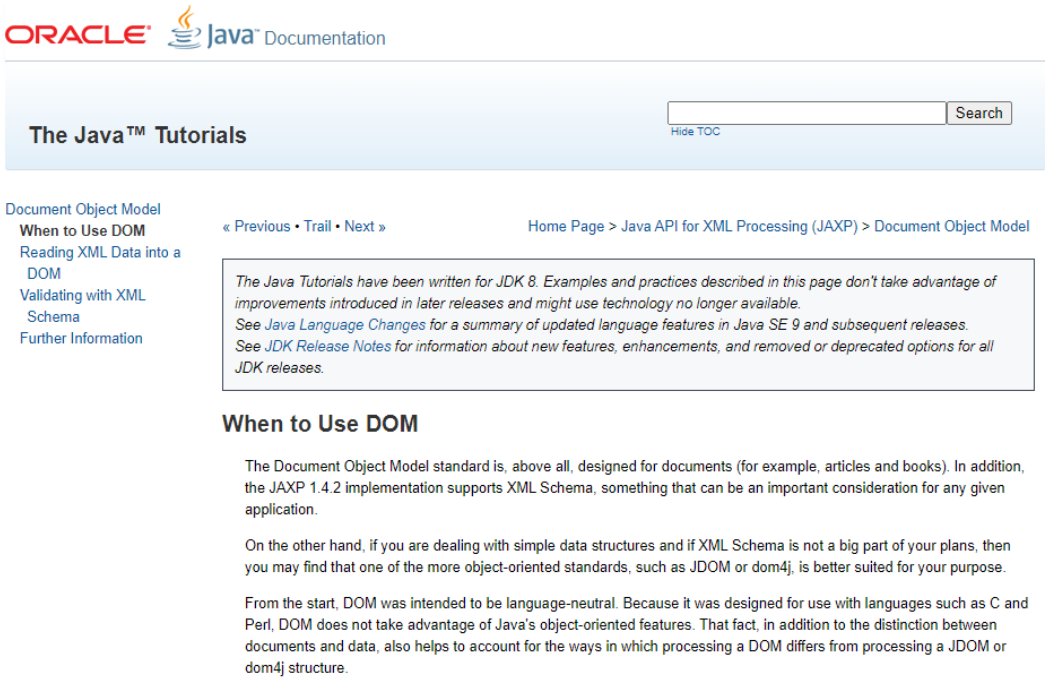
Pretende-se com esta ficha prática que os alunos implementem um parsing de ficheiros XML

Conteúdo

2. DOM

DOM (Document Object Model) foi criado para linguagens de programação não orientadas a objetos tal como o (JDOM).

Utiliza “Mixed-Content Model” que deixa o texto e os elementos podem correr na mesma página sem haver distinção.


ORACLE Java Documentation

The Java™ Tutorials

Search

Hide TOC

Document Object Model
When to Use DOM
Reading XML Data into a DOM
Validating with XML Schema
Further Information

« Previous • Trail • Next »

Home Page > Java API for XML Processing (JAXP) > Document Object Model

The Java Tutorials have been written for JDK 8. Examples and practices described in this page don't take advantage of improvements introduced in later releases and might use technology no longer available.
See Java Language Changes for a summary of updated language features in Java SE 9 and subsequent releases.
See JDK Release Notes for information about new features, enhancements, and removed or deprecated options for all JDK releases.

When to Use DOM

The Document Object Model standard is, above all, designed for documents (for example, articles and books). In addition, the JAXP 1.4.2 implementation supports XML Schema, something that can be an important consideration for any given application.

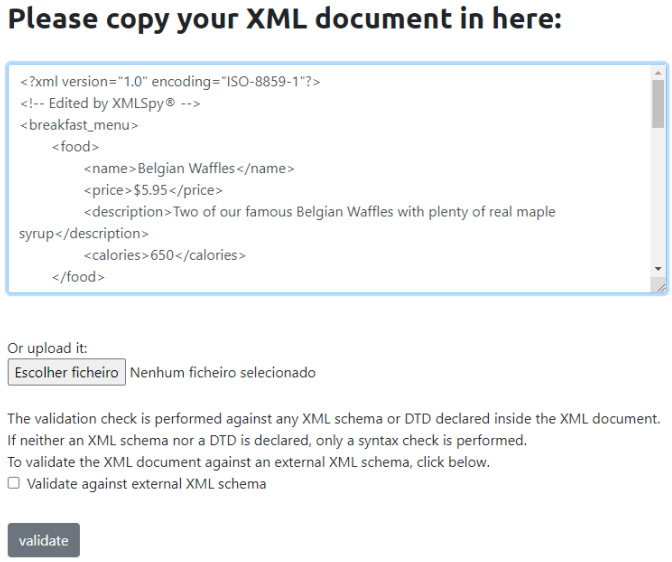
On the other hand, if you are dealing with simple data structures and if XML Schema is not a big part of your plans, then you may find that one of the more object-oriented standards, such as JDOM or dom4j, is better suited for your purpose.

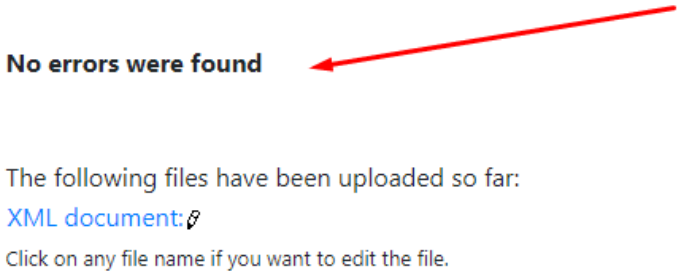
From the start, DOM was intended to be language-neutral. Because it was designed for use with languages such as C and Perl, DOM does not take advantage of Java's object-oriented features. That fact, in addition to the distinction between documents and data, also helps to account for the ways in which processing a DOM differs from processing a JDOM or dom4j structure.

https://docs.oracle.com/javase/tutorial/jaxp/dom/when.html

2.1. DOM

Em primeiro testamos se o documento XML não tem erros de sintax para que possa transferido para o servidor
https://www.xmlvalidation.com/





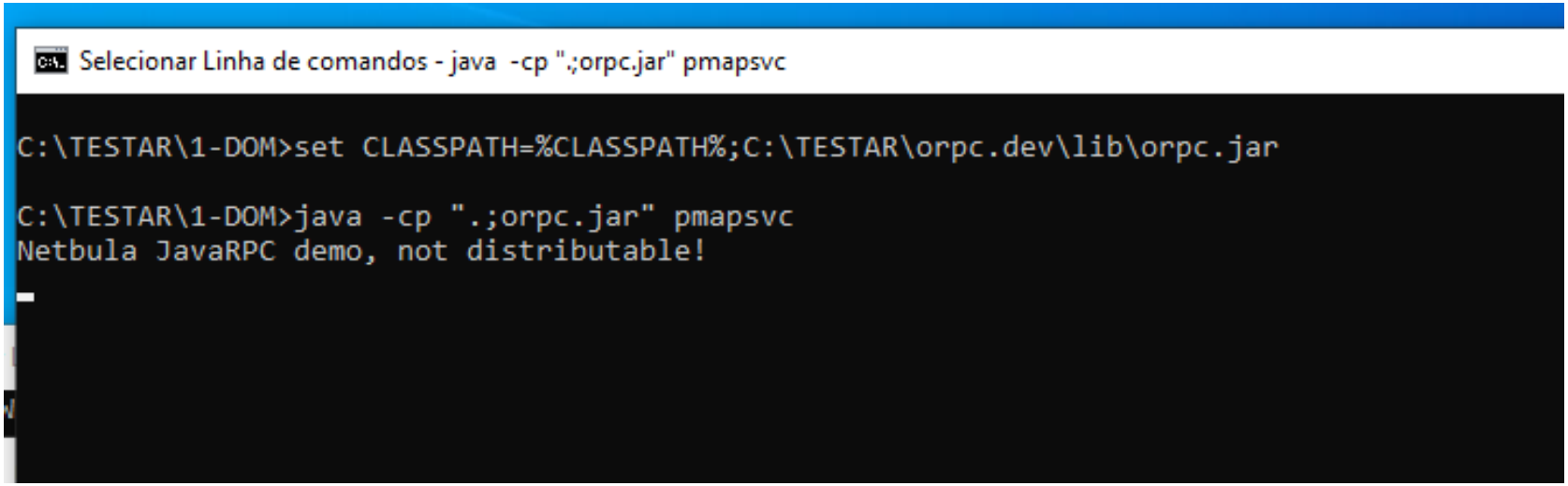
2.2. DOM

Alteramos o código que abre um CMD e executa o DOM do ficheiro anexado tal como abordado no Tutorial.

```
public class FileServer extends putfile_svcb{

    public void sendFile(NFiles in_arg){

            for(int i=0; i<in_arg.files.length; i++) {
        System.out.println("Ficheiro recebido: "+ in_arg.files[i].receivedFilepath()+ " " + in_arg.files[i].byteCount()+ " bytes transferidos");
        System.out.println("Ficheiro gravado: "+ in_arg.files[i].savedFilename());

        try{
        Runtime.getRuntime().exec("cmd.exe /k start java dom/DOMEcho "+ in_arg.files[i].savedFilename());

        }catch (IOException  e)
        {
            e.printStackTrace();
        }
        }
        }
    static public void main(String args[]) {
```

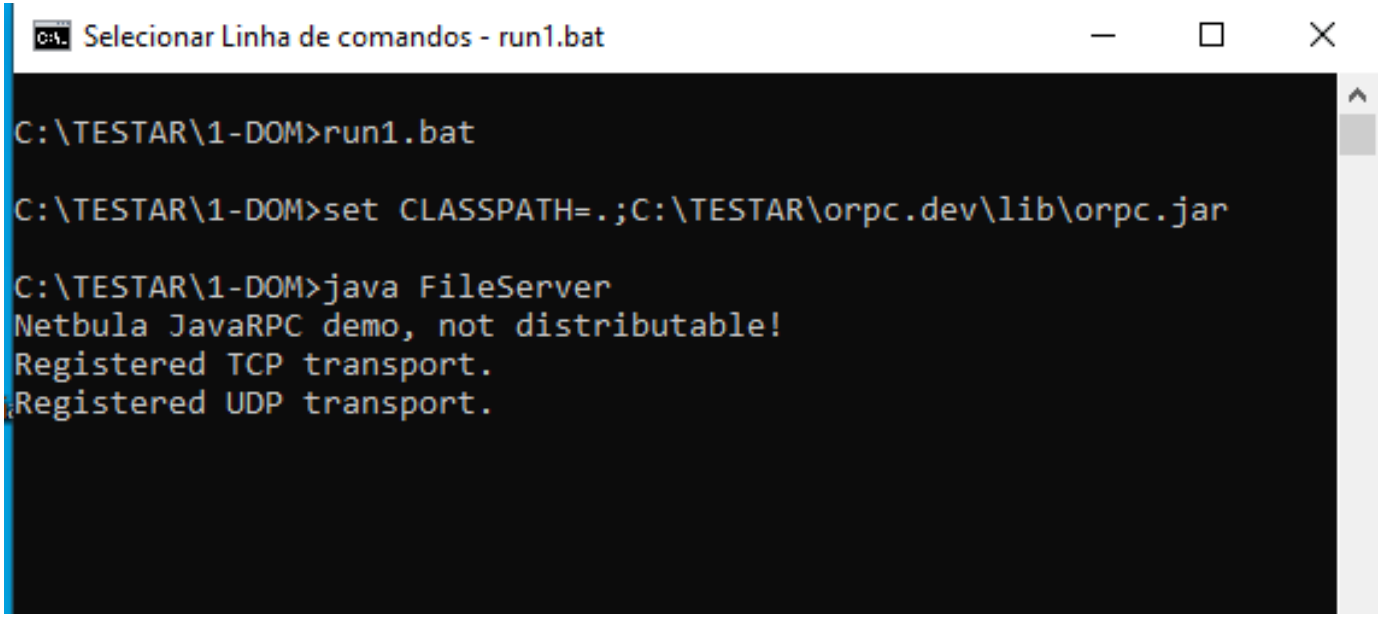
2.3. DOM

Executar o pmapsvc (PortMapper) no primeiro CMD.



2.4. DOM

Executar o servidor, como podemos constatar os pacotes de TCP e UDP estão operacionais.



2.5. DOM

Executar os comandos para podemos usar o compilador do java e o jrpcgen que é uma ferramenta que gera código java para implementar um RPC protocolo.

setup.bat - Bloco de notas

Ficheiro Editar Formatar Ver Ajuda

```
set PATH=%PATH%;c:\TESTAR\orpc.dev\bin\win32
set PATH=%PATH%;"C:\Program Files\Java\jdk1.8.0_301\bin"
set CLASSPATH=.
```

Linha de comandos

```
C:\TESTAR\1-DOM>setup.bat

C:\TESTAR\1-DOM>set PATH=C:\Program Files (x86)\Common Files\Oracle\
Java\javapath;C:\Windows\system32;C:\Windows;C:\Windows\System32\Wbe
m;C:\Windows\System32\WindowsPowerShell\v1.0\;C:\Windows\System32\Op
enSSH\;C:\Program Files\nodejs\;C:\Users\Ruben\AppData\Local\Android
\Sdk\tools\bin;C:\Users\Ruben\AppData\Local\Android\Sdk\platform-too
ls;C:\Users\Ruben\AppData\Local\Android\Sdk\emulator;C:\Program File
s\Git\cmd;C:\Users\Ruben\AppData\Local\Programs\Python\Python310\Scr
ipts\;C:\Users\Ruben\AppData\Local\Programs\Python\Python310\;C:\Use
rs\Ruben\AppData\Local\Microsoft\WindowsApps;C:\Users\Ruben\AppData\
Local\Programs\Microsoft VS Code\bin;C:\Users\Ruben\AppData\Roaming\
npm;C:\Program Files\JetBrains\IntelliJ IDEA Community Edition 2021.
2.2\bin;;C:\Program Files\JetBrains\PyCharm 2021.2.2\bin;;;c:\TESTAR
\orpc.dev\bin\win32;"C:\Program Files\Java\jdk1.8.0_301\bin";c:\TEST
AR\orpc.dev\bin\win32;"C:\Program Files\Java\jdk1.8.0_301\bin";c:\TE
STAR\orpc.dev\bin\win32

C:\TESTAR\1-DOM>set PATH=C:\Program Files (x86)\Common Files\Oracle\
Java\javapath;C:\Windows\system32;C:\Windows;C:\Windows\System32\Wbe
m;C:\Windows\System32\WindowsPowerShell\v1.0\;C:\Windows\System32\Op
enSSH\;C:\Program Files\nodejs\;C:\Users\Ruben\AppData\Local\Android
\Sdk\tools\bin;C:\Users\Ruben\AppData\Local\Android\Sdk\platform-too
ls;C:\Users\Ruben\AppData\Local\Android\Sdk\emulator;C:\Program File
s\Git\cmd;C:\Users\Ruben\AppData\Local\Programs\Python\Python310\Scr
ipts\;C:\Users\Ruben\AppData\Local\Programs\Python\Python310\;C:\Use
```

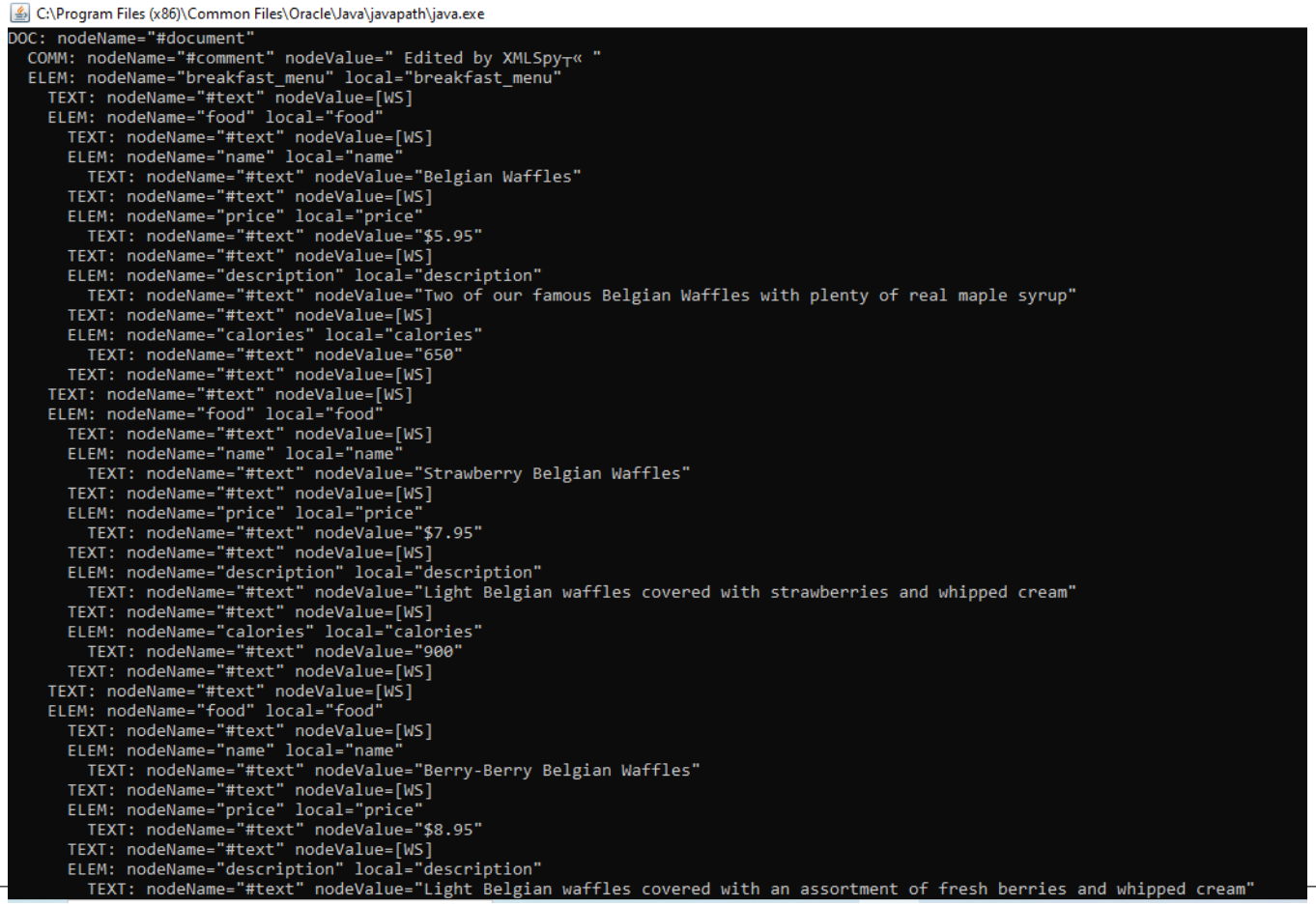
2.6. DOM

Executar o cliente e anexar o ficheiro simple.xml.

```
C:\TESTAR\1-DOM>set CLASSPATH=.;C:\TESTAR\orpc.dev\lib\orpc.jar

C:\TESTAR\1-DOM>java FileClient localhost simple.xml
Netbula JavaRPC demo, not distributable!
Ligado ao servidor  localhost
```

2.7. DOM

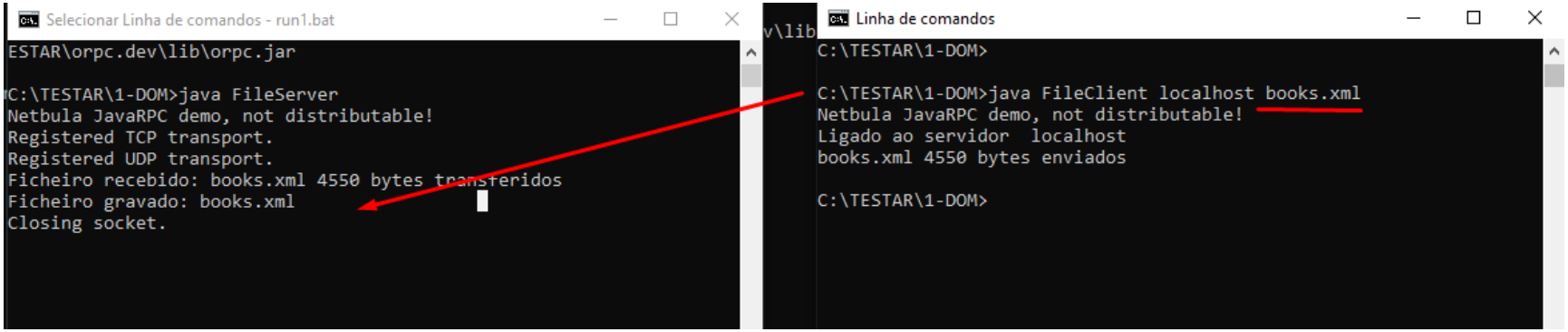
Após executar o cliente foi aberto uma nova janela de CMD com o nosso ficheiro XML no servidor.



2.8. DOM

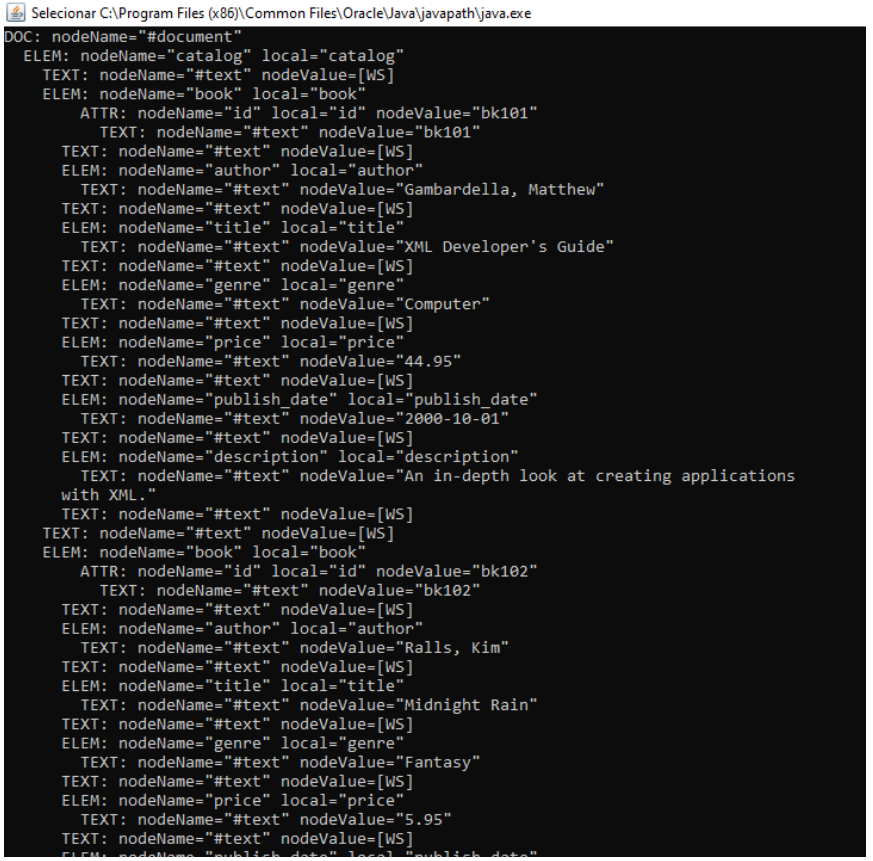
Depois de enviarmos alguns ficheiros XML através do cliente para o servidor conseguimos constatar que o servidor nos dá a resposta dizendo que os ficheiros foram gravados.

```
Ficheiro gravado: simple.xml
Ficheiro gravado: medsamp2014.xml
Closing socket.
```



2.9. DOM

Nesta imagem temos o XML do ficheiro books.xml.



2.10. DOM

Depois de testarmos o ficheiro medsamp.xml percebemos que o ficheiro tinha erros de syntax do qual fomos testar no site: (https://www.xmlvalidation.com/) onde tivemos a resposta que o havia um problema com o ficheiro.

Please copy your XML document in here:

```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE MedlineCitationSet PUBLIC "-//NLM//DTD Medline Citation, 1st January, 2014//EN"
"http://www.nlm.nih.gov/databases/dtd/nlmmedlinecitationset_140101.dtd">
<MedlineCitationSet>
<MedlineCitation Owner="NLM" Status="MEDLINE">
<PMID Version="1">10540283</PMID>
<DateCreated>
<Year>1999</Year>
<Month>12</Month>
<Day>17</Day>
```

Or upload it:

Escolher ficheiro medsamp2014.xml

The validation check is performed against any XML schema or DTD declared inside the XML document.
If neither an XML schema nor a DTD is declared, only a syntax check is performed.
To validate the XML document against an external XML schema, click below.

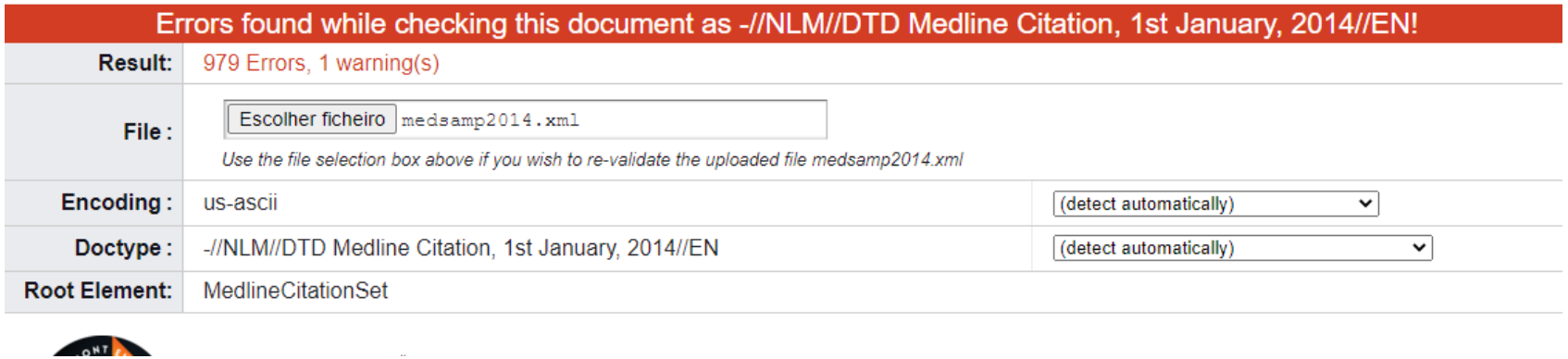
☐ Validate against external XML schema

validate

A program error occured. Sorry.
We will analyze this problem soon.

2.11. DOM

Testamos o mesmo ficheiro no site (https://validator.w3.org/) onde obtivemos os seguintes erros.



2.12. DOM

Breve descrição os erros:

Validation Output: 979 Errors

Line 2, Column 162: **DTD did not contain element declaration for document type name**

```
…4//EN" "http://www.nlm.nih.gov/databases/dtd/nlmmedlinecitationset_140101.dtd">
```

A DOCTYPE declares the version of the language used, as well as what the root (top) element of your document will be. For example, if the top element of your document is <html>, the DOCTYPE declaration will look like: "<!DOCTYPE html".

In most cases, it is safer not to type or edit the DOCTYPE declaration at all, and preferable to let a tool include it, or copy and paste it from a trusted list of DTDs.

Line 3, Column 20: **element "MedlineCitationSet" undefined**

```
<MedlineCitationSet>
```

You have used the element named above in your document, but the document type you are using does not define an element of that name. This error is often caused by:

- incorrect use of the "Strict" document type with a document that uses frames (e.g. you must use the "Frameset" document type to get the "<frameset>" element),
- by using vendor proprietary extensions such as "<spacer>" or "<marquee>" (this is usually fixed by using CSS to achieve the desired effect instead).
- by using upper-case tags in XHTML (in XHTML attributes and elements must be all lower-case).

Line 4, Column 24: **there is no attribute "Owner"**

```
<MedlineCitation Owner="NLM" Status="MEDLINE">
```

You have used the attribute named above in your document, but the document type you are using does not support that attribute for this element. This error is often caused by incorrect use of the "Strict" document type with a document that uses frames (e.g. you must use the "Transitional" document type to get the "target" attribute), or by using vendor proprietary extensions such as "marginheight" (this is usually fixed by using CSS to achieve the desired effect instead).

This error may also result if the element itself is not supported in the document type you are using, as an undefined element will have no supported attributes; in this case, see the element-undefined

2.13. DOM

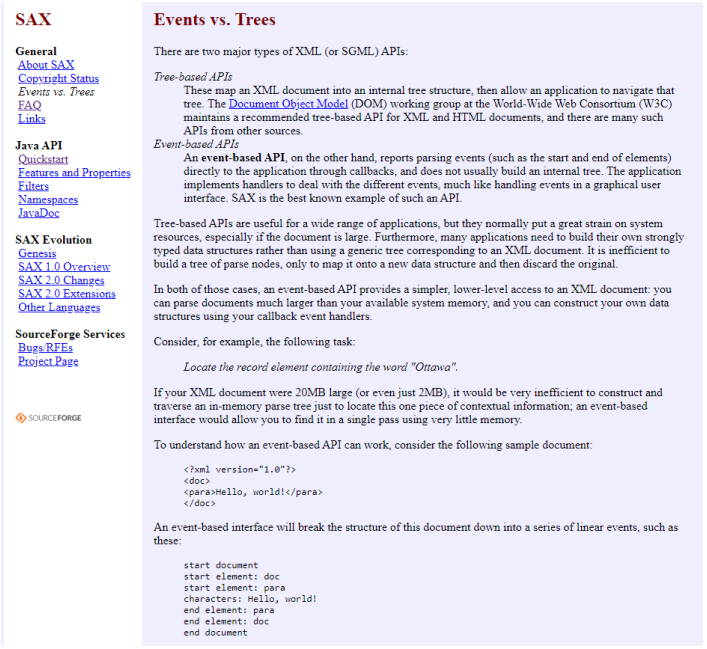
Após corrigir os erros enviamos para o servidor o ficheiro medsamp2014.xml.

```
DOC: nodeName="#document"
  ELEM: nodeName="MedlineCitationSet" local="MedlineCitationSet"
    TEXT: nodeName="#text" nodeValue=[WS]
    ELEM: nodeName="MedlineCitation" local="MedlineCitation"
        ATTR: nodeName="Owner" local="Owner" nodeValue="NLM"
          TEXT: nodeName="#text" nodeValue="NLM"
        ATTR: nodeName="Status" local="Status" nodeValue="MEDLINE"
          TEXT: nodeName="#text" nodeValue="MEDLINE"
      TEXT: nodeName="#text" nodeValue=[WS]
      ELEM: nodeName="PMID" local="PMID"
          ATTR: nodeName="Version" local="Version" nodeValue="1"
            TEXT: nodeName="#text" nodeValue="1"
        TEXT: nodeName="#text" nodeValue="10540283"
      TEXT: nodeName="#text" nodeValue=[WS]
      ELEM: nodeName="DateCreated" local="DateCreated"
        TEXT: nodeName="#text" nodeValue=[WS]
        ELEM: nodeName="Year" local="Year"
          TEXT: nodeName="#text" nodeValue="1999"
        TEXT: nodeName="#text" nodeValue=[WS]
        ELEM: nodeName="Month" local="Month"
          TEXT: nodeName="#text" nodeValue="12"
        TEXT: nodeName="#text" nodeValue=[WS]
        ELEM: nodeName="Day" local="Day"
          TEXT: nodeName="#text" nodeValue="17"
        TEXT: nodeName="#text" nodeValue=[WS]
      TEXT: nodeName="#text" nodeValue=[WS]
      ELEM: nodeName="DateCompleted" local="DateCompleted"
        TEXT: nodeName="#text" nodeValue=[WS]
        ELEM: nodeName="Year" local="Year"
          TEXT: nodeName="#text" nodeValue="1999"
        TEXT: nodeName="#text" nodeValue=[WS]
        ELEM: nodeName="Month" local="Month"
          TEXT: nodeName="#text" nodeValue="12"
        TEXT: nodeName="#text" nodeValue=[WS]
        ELEM: nodeName="Day" local="Day"
          TEXT: nodeName="#text" nodeValue="17"
        TEXT: nodeName="#text" nodeValue=[WS]
      TEXT: nodeName="#text" nodeValue=[WS]
      ELEM: nodeName="DateRevised" local="DateRevised"
        TEXT: nodeName="#text" nodeValue=[WS]
        ELEM: nodeName="Year" local="Year"
          TEXT: nodeName="#text" nodeValue="2006"
        TEXT: nodeName="#text" nodeValue=[WS]
        ELEM: nodeName="Month" local="Month"
```

```
C:\TESTAR\1-DOM>java FileClient localhost medsamp2014.xml
Netbula JavaRPC demo, not distributable!
Ligado ao servidor  localhost
medsamp2014.xml 63732 bytes enviados
```

3. SAX

SAX (Simple API for XML) foi o primeiro a adotar o XML para java, é uma API baseada em eventos.
Ao contrário do DOM que é "Tree-based API" que mapeia um XML documento para uma estrutura em árvore. Este tipo é mais recomendado para documentos tipo HTML.



SAX

General
About SAX
Copyright Status
Events vs. Trees
FAQ
Links

Java API
Quickstart
Features and Properties
Filters
Namespaces
JavaDoc

SAX Evolution
Genesis
SAX 1.0 Overview
SAX 2.0 Changes
SAX 2.0 Extensions
Other Languages

SourceForge Services
Bugs/RFEs
Project Page

SOURCEFORGE

Events vs. Trees

There are two major types of XML (or SGML) APIs:

Tree-based APIs
These map an XML document into an internal tree structure, then allow an application to navigate that tree. The Document Object Model (DOM) working group at the World-Wide Web Consortium (W3C) maintains a recommended tree-based API for XML and HTML documents, and there are many such APIs from other sources.

Event-based APIs
An **event-based API**, on the other hand, reports parsing events (such as the start and end of elements) directly to the application through callbacks, and does not usually build an internal tree. The application implements handlers to deal with the different events, much like handling events in a graphical user interface. SAX is the best known example of such an API.

Tree-based APIs are useful for a wide range of applications, but they normally put a great strain on system resources, especially if the document is large. Furthermore, many applications need to build their own strongly typed data structures rather than using a generic tree corresponding to an XML document. It is inefficient to build a tree of parse nodes, only to map it onto a new data structure and then discard the original.

In both of those cases, an event-based API provides a simpler, lower-level access to an XML document: you can parse documents much larger than your available system memory, and you can construct your own data structures using your callback event handlers.

Consider, for example, the following task:

> *Locate the record element containing the word "Ottawa".*

If your XML document were 20MB large (or even just 2MB), it would be very inefficient to construct and traverse an in-memory parse tree just to locate this one piece of contextual information; an event-based interface would allow you to find it in a single pass using very little memory.

To understand how an event-based API can work, consider the following sample document:

```
<?xml version="1.0"?>
<doc>
<para>Hello, world!</para>
</doc>
```

An event-based interface will break the structure of this document down into a series of linear events, such as these:

```
start document
start element: doc
start element: para
characters: Hello, world!
end element: para
end element: doc
end document
```

http://www.saxproject.org/event.html

3.1. SAX

Mais uma vez alteramos o código que abre um CMD e executa o JDOM do ficheiro anexado tal como abordado no Tutorial.

```
try{
    Runtime.getRuntime().exec("cmd.exe /k start java sax/SAXLocalNameCount "+
            in_arg.files[i].savedFilename());
```

3.2. SAX

Mais uma vez executar o pmapsvc (PortMapper) no primeiro CMD.

```
C:\TESTAR\2-SAX>set CLASSPATH=%CLASSPATH%;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc
.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar

C:\TESTAR\2-SAX>java -cp ".;orpc.jar" pmapsvc
Netbula JavaRPC demo, not distributable!
```

3.3. SAX

Executar o servidor no segundo CMD.

```
C:\TESTAR\2-SAX>set CLASSPATH=.;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\T
ESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar

C:\TESTAR\2-SAX>java FileServer
Netbula JavaRPC demo, not distributable!
Registered TCP transport.
Registered UDP transport.
```

3.4. SAX

Ao executar o programa SAXLocalNameCount retorna uma listagem das palavras e o numero de vezes que são encontradas no ficheiro books.xml

```
C:\TESTAR\2-SAX>java sax/SAXLocalNameCount "C:\TESTAR\2-SAX\books.xml"
Local Name "genre" occurs 12 times
Local Name "catalog" occurs 1 times
Local Name "book" occurs 12 times
Local Name "author" occurs 12 times
Local Name "description" occurs 12 times
Local Name "price" occurs 12 times
Local Name "publish_date" occurs 12 times
Local Name "title" occurs 12 times
```

3.5. SAX

Mais uma vez ao executar o programa SAXLocalNameCount retorna uma listagem das palavras e o numero de vezes que são encontradas nesta vez no ficheiro note.xml

```
C:\TESTAR\2-SAX>java sax/SAXLocalNameCount "C:\TESTAR\2-SAX\note.xml"
Local Name "body" occurs 1 times
Local Name "note" occurs 1 times
Local Name "heading" occurs 1 times
Local Name "from" occurs 1 times
Local Name "to" occurs 1 times
```

3.6. SAX

Voltamos a constatar que o ficheiro tinha erros de Syntax

*The maximum size limit for file upload is 2 megabytes.

- XML Document Structures Must Start And End Within The Same Entity., Line '3146', Column '23'.
- XML Document Structures Must Start And End Within The Same Entity.

Close

3.7. SAX

Os erros foram corrigidos ao fechar algumas tags.

```
3136                <QualifierName MajorTopicYN="Y">immunology</QualifierName>
3137            </MeshHeading>
3138            <MeshHeading>
3139                <DescriptorName MajorTopicYN="N">Glucuronidase</DescriptorName>
3140                <QualifierName MajorTopicYN="N">secretion</QualifierName>
3141            </MeshHeading>
3142            <MeshHeading>
3143                <DescriptorName MajorTopicYN="N">Histamine</DescriptorName>
3144                <QualifierName MajorTopicYN="Y">analogs &amp;amp; derivatives</QualifierName>
3145            </MeshHeading>
3146        </MeshHeadingList>
3147    </MedlineCitation>
3148 </MedlineCitationSet>
3149
3150
```

3.8. SAX

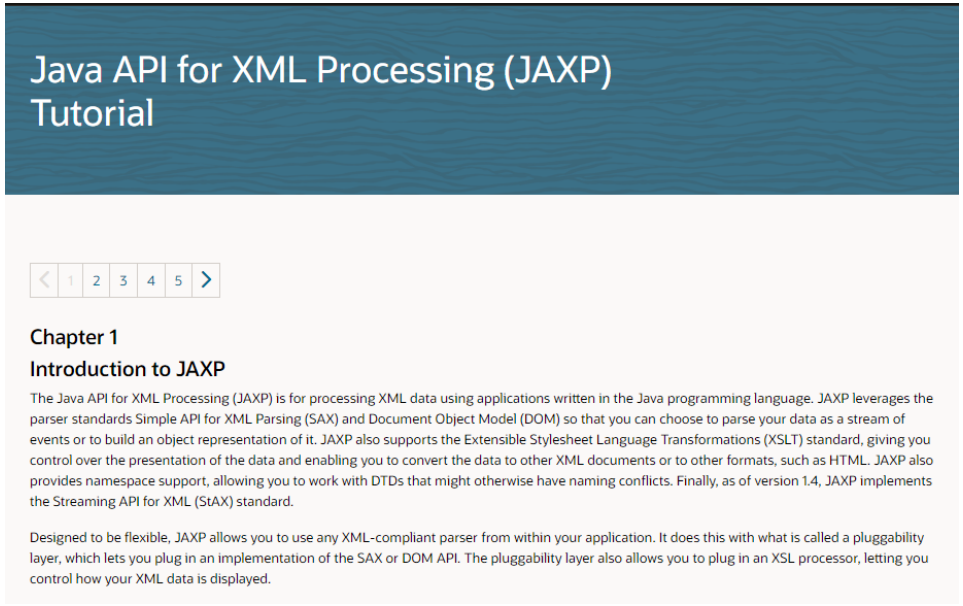
Depois da resolução dos problemas corremos o programa e obtivemos a seguinte resposta.

```
C:\TESTAR\2-SAX>java sax/SAXLocalNameCount "C:\TESTAR\2-SAX\medsamp2013.xml"
Local Name "RefSource" occurs 9 times
Local Name "Grant" occurs 6 times
Local Name "NameOfSubstance" occurs 47 times
Local Name "AbstractText" occurs 18 times
Local Name "AccessionNumber" occurs 4 times
Local Name "AuthorList" occurs 11 times
Local Name "MeshHeadingList" occurs 10 times
Local Name "ForeName" occurs 328 times
Local Name "ArticleTitle" occurs 11 times
Local Name "MedlineCitationSet" occurs 1 times
Local Name "Volume" occurs 10 times
Local Name "ChemicalList" occurs 7 times
Local Name "Day" occurs 36 times
Local Name "Affiliation" occurs 11 times
Local Name "ISOAbbreviation" occurs 11 times
Local Name "Note" occurs 1 times
Local Name "Acronym" occurs 5 times
Local Name "CopyrightInformation" occurs 2 times
Local Name "DescriptorName" occurs 151 times
Local Name "PublicationType" occurs 25 times
Local Name "MedlineJournalInfo" occurs 11 times
Local Name "MedlineTA" occurs 11 times
Local Name "DateCompleted" occurs 10 times
Local Name "OtherID" occurs 2 times
Local Name "Title" occurs 11 times
Local Name "PMID" occurs 18 times
Local Name "MeshHeading" occurs 151 times
Local Name "CommentsCorrectionsList" occurs 3 times
Local Name "DataBankName" occurs 2 times
Local Name "ISSN" occurs 11 times
Local Name "PublicationTypeList" occurs 11 times
Local Name "JournalIssue" occurs 11 times
Local Name "Agency" occurs 6 times
Local Name "PubDate" occurs 11 times
Local Name "Pagination" occurs 11 times
Local Name "Article" occurs 11 times
Local Name "DataBankList" occurs 1 times
Local Name "KeywordList" occurs 1 times
Local Name "GrantID" occurs 6 times
```

```
Local Name "Chemical" occurs 47 times
Local Name "ISSNLinking" occurs 11 times
Local Name "MedlinePgn" occurs 11 times
Local Name "CitationSubset" occurs 13 times
Local Name "Author" occurs 329 times
Local Name "AccessionNumberList" occurs 2 times
Local Name "QualifierName" occurs 83 times
Local Name "Issue" occurs 10 times
Local Name "MedlineCitation" occurs 11 times
Local Name "Country" occurs 17 times
Local Name "DataBank" occurs 2 times
Local Name "GrantList" occurs 2 times
Local Name "Journal" occurs 11 times
Local Name "CommentsCorrections" occurs 9 times
Local Name "NumberOfReferences" occurs 1 times
Local Name "GeneralNote" occurs 1 times
Local Name "OtherAbstract" occurs 1 times
Local Name "MedlineDate" occurs 1 times
Local Name "Month" occurs 41 times
Local Name "ArticleDate" occurs 1 times
Local Name "Initials" occurs 328 times
Local Name "DateCreated" occurs 11 times
Local Name "Language" occurs 11 times
Local Name "CollectiveName" occurs 1 times
Local Name "NlmUniqueID" occurs 11 times
Local Name "Keyword" occurs 26 times
Local Name "RegistryNumber" occurs 47 times
Local Name "LastName" occurs 328 times
Local Name "Abstract" occurs 11 times
Local Name "DateRevised" occurs 10 times
Local Name "Year" occurs 42 times
```

4. JAXP

JAXP (Java API for XML Processing) processa XML data usando aplicações em Java, podemos escolher se queremos fazer parsing da data como stream de eventos ou como object based.



https://www.oracle.com/java/technologies/jaxp-introduction.html

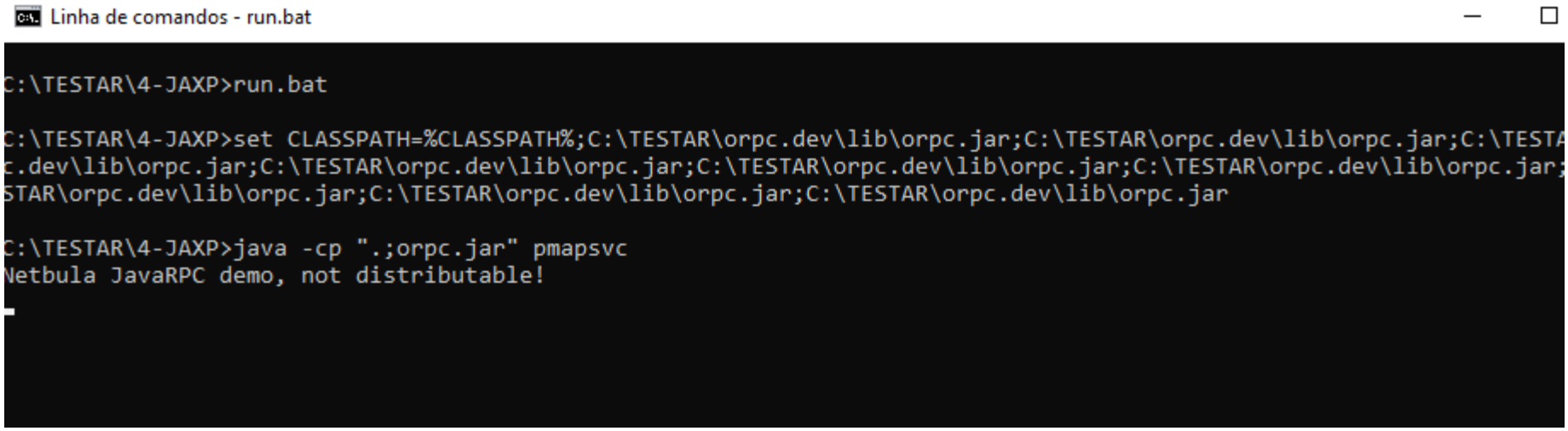
4.1. JAXP

Alteração do Runtime para “stax/cursor/CursorParse” para correr o JAXP

```
try{
    Runtime.getRuntime().exec( "cmd.exe /c start java stax/cursor/CursorParse "+in_arg.files[i].savedFilename() );
```

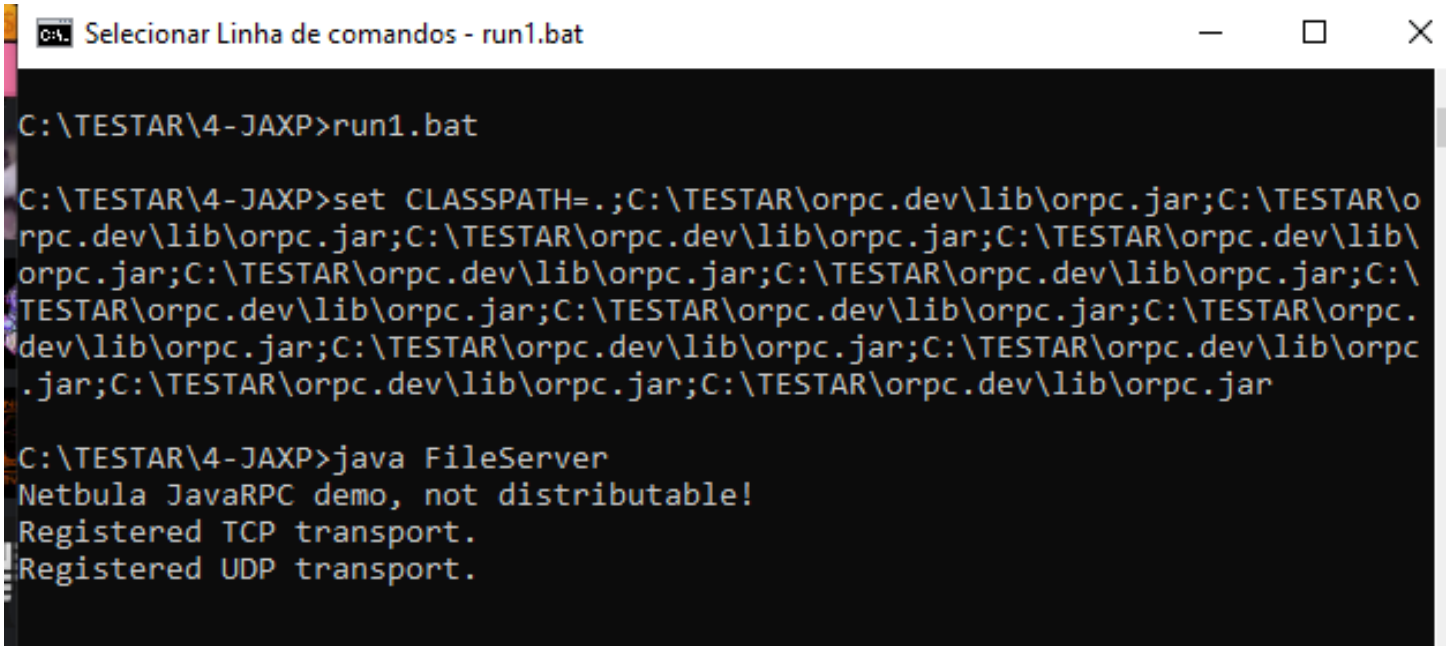
4.2. JAXP

Mais uma vez executar o pmapsvc (PortMapper).



4.3. JAXP

Executar o servidor no segundo CMD.



4.4. JAXP

Executar o cliente e em anexo o ficheiro de books.xml que é enviado para o servidor.

```
C:\TESTAR\4-JAXP>set CLASSPATH=.;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TE
STAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\o
rpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.de
v\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\
orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.j
ar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar

C:\TESTAR\4-JAXP>java FileClient localhost books.xml
Netbula JavaRPC demo, not distributable!
Ligado ao servidor  localhost
books.xml 4550 bytes enviados
```

4.5. JAXP

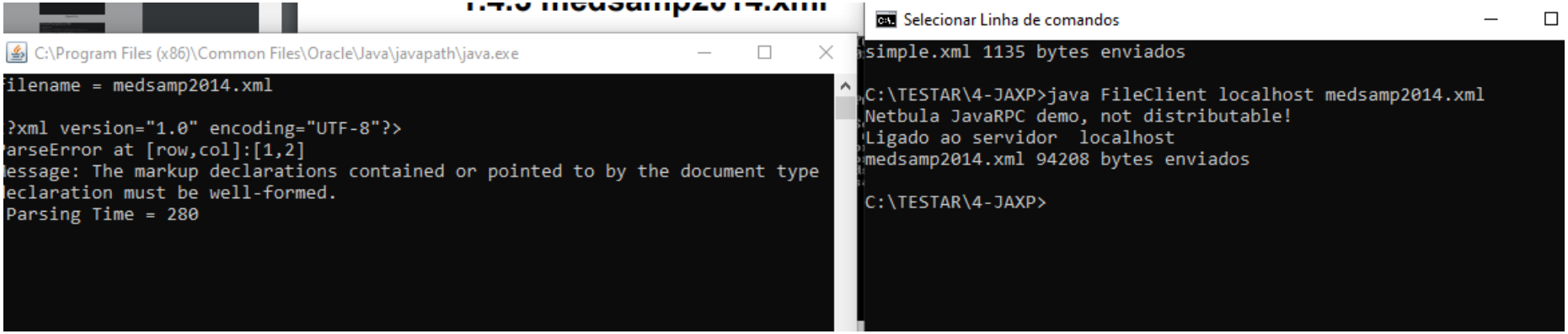
Após enviar o ficheiro xml para o servidor é aberto um novo terminal com o xml recebido.

Selecionar C:\Program Files (x86)\Common Files\Oracle\Java\javapath\java.exe

```
FACTORY: com.sun.xml.internal.stream.XMLInputFactoryImpl@5c647e05
filename = books.xml

<?xml version="1.0" encoding="null"?>
<catalog>
   <book id="bk101">
      <author>Gambardella, Matthew</author>
      <title>XML Developer's Guide</title>
      <genre>Computer</genre>
      <price>44.95</price>
      <publish_date>2000-10-01</publish_date>
      <description>An in-depth look at creating applications
      with XML.</description>
   </book>
   <book id="bk102">
      <author>Ralls, Kim</author>
      <title>Midnight Rain</title>
      <genre>Fantasy</genre>
      <price>5.95</price>
      <publish_date>2000-12-16</publish_date>
      <description>A former architect battles corporate zombies,
      an evil sorceress, and her own childhood to become queen
      of the world.</description>
   </book>
   <book id="bk103">
      <author>Corets, Eva</author>
      <title>Maeve Ascendant</title>
      <genre>Fantasy</genre>
      <price>5.95</price>
      <publish_date>2000-11-17</publish_date>
      <description>After the collapse of a nanotechnology
      society in England, the young survivors lay the
      foundation for a new society.</description>
   </book>
   <book id="bk104">
      <author>Corets, Eva</author>
      <title>Oberon's Legacy</title>
      <genre>Fantasy</genre>
      <price>5.95</price>
      <publish_date>2001-03-10</publish_date>
      <description>In post-apocalypse England, the mysterious
      agent known only as Oberon helps to create a new life
      for the inhabitants of London. Sequel to Maeve
```

4.6. JAXP

Após enviar o ficheiro XML para o servidor foi nos retornado um erro no entanto o cliente conseguiu enviar o documento para o servidor independentemente de ter erros ou não.



4.7. JAXP

O erro no ficheiro medsamp2014.xml era unicamente fechar uma tag.

```
                    <Initials>KW</Initials>
                </Author>
            </AuthorList>
            <Language>eng</Language>
            <PublicationTypeList>
                <PublicationType>Journal Article</PublicationType>
            </PublicationTypeList>
        </Article>
        <MedlineJournalInfo>
            <Country>UNITED STATES</Country>
            <MedlineTA>Phys Rev Lett</MedlineTA>
            <NlmUniqueID>0401141</NlmUniqueID>
            <ISSNLinking>0031-9007</ISSNLinking>
        </MedlineJournalInfo>
        <CitationSubset>IM</CitationSubset>
    </MedlineCitation>1
</MedlineCitationSet>
```

4.8. JAXP

Após corrigir o erro conseguimos correr o ficheiro sem problema.

```
FACTORY: com.sun.xml.internal.stream.XMLInputFactoryImpl@5c647e05
filename = medsamp2014.xml

<?xml version="1.0" encoding="UTF-8"?>
<MedlineCitationSet>
    <MedlineCitation Owner="NLM" Status="MEDLINE">
        <PMID Version="1">10540283</PMID>
        <DateCreated>
            <Year>1999</Year>
            <Month>12</Month>
            <Day>17</Day>
        </DateCreated>
        <DateCompleted>
            <Year>1999</Year>
            <Month>12</Month>
            <Day>17</Day>
        </DateCompleted>
        <DateRevised>
            <Year>2006</Year>
            <Month>11</Month>
            <Day>15</Day>
        </DateRevised>
        <Article PubModel="Print">
            <Journal>
                <ISSN IssnType="Print">0950-382X</ISSN>
                <JournalIssue CitedMedium="Print">
                    <Volume>34</Volume>
                    <Issue>1</Issue>
                    <PubDate>
                        <Year>1999</Year>
                        <Month>Oct</Month>
                    </PubDate>
                </JournalIssue>
                <Title>Molecular microbiology</Title>
                <ISOAbbreviation>Mol. Microbiol.</ISOAbbreviation>
            </Journal>
            <ArticleTitle>Transcription regulation of the nir gene cluster encoding nitrite reductase of Paracoccus denitrificans involves NNR and NirI, a novel type of
 membrane protein.</ArticleTitle>
            <Pagination>
                <MedlinePgn>24-36</MedlinePgn>
            </Pagination>
            <Abstract>
                <AbstractText>The nirIX gene cluster of Paracoccus denitrificans is located between the nir and nor gene clusters encoding nitrite and nitric oxide redu
```

5. JAXB

JAXB (Java Architecture for XML Binding) providencia uma API e ferramentas para automatizar o mapeamento entre o XML e os objetos em java.

JAXB

The Java™ Architecture for XML Binding (JAXB) provides an API and tools that automate the mapping between XML documents and Java objects.

The JAXB framework enables developers to perform the following operations:

- Unmarshal XML content into a Java representation
- Access and update the Java representation
- Marshal the Java representation of the XML content into XML content

JAXB gives Java developers an efficient and standard way of mapping between XML and Java code. Java developers using JAXB are more productive because they can write less code themselves and do not have to be experts in XML. JAXB makes it easier for developers to extend their applications with XML and Web Services technologies.

https://github.com/javaee/jaxb-v2

5.1 JAXB

Como já feito nos exemplos anteriores vamos alterar a instrução do Runtime para executar JAXB.

```
try{
 Runtime.getRuntime().exec( "cmd.exe /c start java Main "+in_arg.files[i].savedFilename() );


}
```

5.2 JAXB

Executar novamente o pmapsvc (PortMapper)

```
C:\TESTAR\5-JAXB\samples\unmarshal-read>run.bat

C:\TESTAR\5-JAXB\samples\unmarshal-read>set CLASSPATH=%CLASSPATH%;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib
\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc
.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TES
TAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;E:\3 ano\IS\1314-EI-IS-Ficha-3-javarpc\orpc.dev\lib\orpc.jar;E
:\3 ano\IS\1314-EI-IS-Ficha-3-javarpc\orpc.dev\lib\orpc.jar

C:\TESTAR\5-JAXB\samples\unmarshal-read>java -cp ".;orpc.jar" pmapsvc
Netbula JavaRPC demo, not distributable!
```

5.3 JAXB

Executar o servidor que fica a espera de pedidos.

```
C:\TESTAR\5-JAXB\samples\unmarshal-read>run1.bat

C:\TESTAR\5-JAXB\samples\unmarshal-read>set CLASSPATH=.;C:\TESTAR\orpc.dev\l
ib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;
C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\or
pc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\o
rpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\T
ESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.d
ev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.
jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;E:\3 ano
\IS\1314-EI-IS-Ficha-3-javarpc\orpc.dev\lib\orpc.jar;E:\3 ano\IS\1314-EI-IS-
Ficha-3-javarpc\orpc.dev\lib\orpc.jar

C:\TESTAR\5-JAXB\samples\unmarshal-read>java FileServer
Netbula JavaRPC demo, not distributable!
Registered TCP transport.
Registered UDP transport.
```

5.4 JAXB

Exemplo do XML que o cliente vai enviar para o servidor.

po.xml

C: > TESTAR > 5-JAXB > samples > unmarshal-read > po.xml > purchaseOrder > items > item

```
 only if the new code is made subject to such option by the copyright
 holder.
-->

<purchaseOrder orderDate="1999-10-20">
  <shipTo country="US">
    <name>Alice Smith</name>
    <street>123 Maple Street</street>
    <city>Cambridge</city>
    <state>MA</state>
    <zip>12345</zip>
  </shipTo>
  <billTo country="US">
    <name>Robert Smith</name>
    <street>8 Oak Avenue</street>
    <city>Cambridge</city>
    <state>MA</state>
    <zip>12345</zip>
  </billTo>
  <items>
    <item partNum="242-NO">
      <productName>Nosferatu - Special Edition (1929)</productName>
      <quantity>5</quantity>
      <USPrice>19.99</USPrice>
    </item>
    <item partNum="242-MU">
      <productName>The Mummy (1959)</productName>
      <quantity>3</quantity>
      <USPrice>19.98</USPrice>
    </item>
    <item partNum="242-GZ">
      <productName>Godzilla and Mothra: Battle for Earth/Godzilla vs. King Ghidora</productName>
      <quantity>3</quantity>
      <USPrice>27.95</USPrice>
    </item>
  </items>
</purchaseOrder>
```

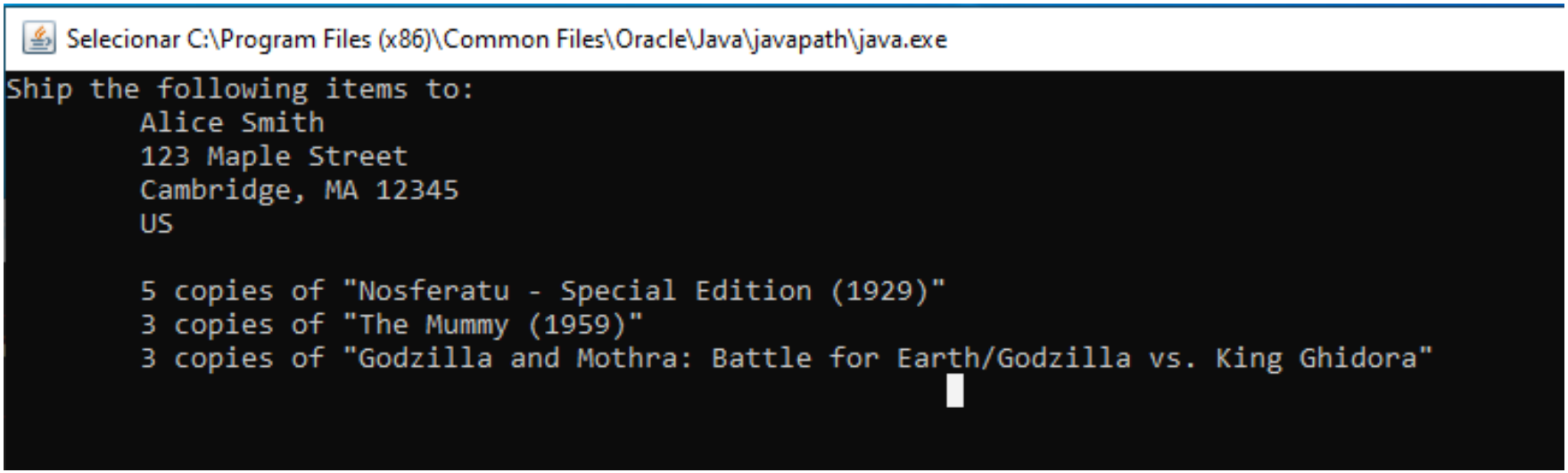
5.5 JAXB

Exatamente como nos exemplos anteriores vamos executar o cliente que vai enviar para o servidor.

```
C:\TESTAR\5-JAXB\samples\unmarshal-read>run2.bat

C:\TESTAR\5-JAXB\samples\unmarshal-read>set CLASSPATH=.;C:\TESTAR\orpc
.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\l
ib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orp
c.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;
C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TES
TAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\or
pc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev
\lib\orpc.jar;E:\3 ano\IS\1314-EI-IS-Ficha-3-javarpc\orpc.dev\lib\orpc
.jar

C:\TESTAR\5-JAXB\samples\unmarshal-read>java FileClient localhost po.x
ml
Netbula JavaRPC demo, not distributable!
Ligado ao servidor  localhost
po.xml 2835 bytes enviados
```

5.6 JAXB

Como podemos verificar o JAXB organizou a nossa informação.



6. JDOM

Execução: Inicialmente ligamos o PortMapper e de seguida ligamos o servidor, nesta implementação em concreto, apesar de o servidor RPC funcionar corretamente, não fazia nenhuma ação o JDOM.

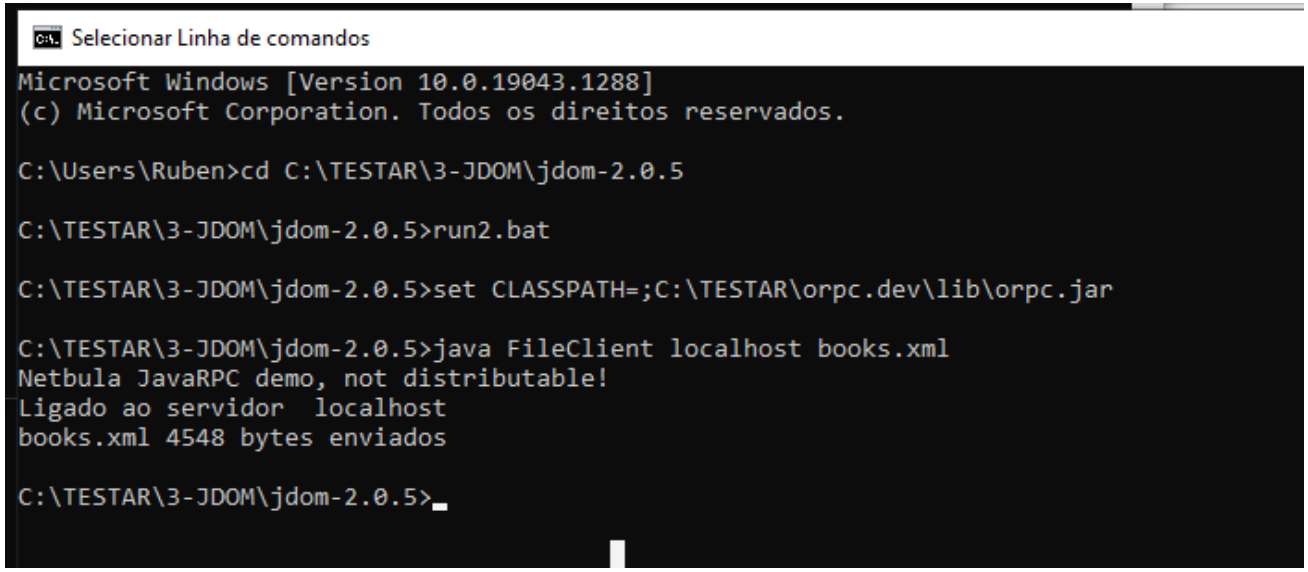
```
try{
    Runtime.getRuntime().exec("cmd.exe /k start java PlayWithJDom "+  in_arg.files[i].savedFilename());
```

```
C:\TESTAR\3-JDOM\jdom-2.0.5>run.bat

C:\TESTAR\3-JDOM\jdom-2.0.5>set CLASSPATH=%CLASSPATH%;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:
\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orp
c.jar;C:\TESTAR\orpc.dev\lib\orpc.jar

C:\TESTAR\3-JDOM\jdom-2.0.5>java -cp ".;orpc.jar" pmapsvc
Netbula JavaRPC demo, not distributable!
```

```
Linha de comandos - run1.bat

C:\TESTAR\3-JDOM\jdom-2.0.5>run1.bat

C:\TESTAR\3-JDOM\jdom-2.0.5>set CLASSPATH=.;C:\TESTAR\orpc.dev\lib\orpc.j
C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR
pc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\lib\orpc.jar;C:\TESTAR\orpc.dev\li
rpc.jar

C:\TESTAR\3-JDOM\jdom-2.0.5>java FileServer
Netbula JavaRPC demo, not distributable!
Registered TCP transport.
Registered UDP transport.
```

6. JDOM

Execução: Depois de ter o servidor ligado , conectamos o cliente escolhendo o ficheiro a ser enviado para o Parser **JDOM**, mas não funcionou.



6. JDOM

Execução: Tentamos correr diretamente o PlayWithDom escolhendo o ficheiro xml a ser efectuado o parsing e descobri mos que estava a dar erro e esse era o motivo pelo o qual o servidor não abria nenhuma janela com o ficheiro em parsing, ainda tentamos aceder aos ficheiros base do JDOM do github: https://github.com/hunterhacker/jdom e mesmo assim não funcionou

```
C:\TESTAR\3-JDOM\jdom-2.0.5>java PlayWithJDom books.xml
Exception in thread "main" java.lang.NoClassDefFoundError: org/jdom2/input/sax/DefaultSAXHandlerFactory$DefaultSAXHandle
r
        at org.jdom2.input.SAXBuilder.<clinit>(SAXBuilder.java:146)
        at PlayWithJDom.parseAndPlay(PlayWithJDom.java:34)
        at PlayWithJDom.main(PlayWithJDom.java:71)
Caused by: java.lang.ClassNotFoundException: org.jdom2.input.sax.DefaultSAXHandlerFactory$DefaultSAXHandler
        at java.net.URLClassLoader.findClass(Unknown Source)
        at java.lang.ClassLoader.loadClass(Unknown Source)
        at sun.misc.Launcher$AppClassLoader.loadClass(Unknown Source)
        at java.lang.ClassLoader.loadClass(Unknown Source)
        ... 3 more

C:\TESTAR\3-JDOM\jdom-2.0.5>
```

6. Conclusão

Concluimos que este trabalho foi muito bom para consolidar os conhecimentos em parsing de XML, com este trabalho aprendemos e executamos diferentes formas de fazer o parsing. Ao longo do trabalho foram aparecendo alguns erros mas sempre superados com destreza, em relação aos ficheiros xml corrompidos nós quando fizemos os tutoriais não reparamos que já tinham mais ficheiros de apoio(Moodle) com as devidas correções dos ficheiros XML e por isso ao longo do trabalho fomos corrigindo os diferentes erros e mostrando, uma experiência que consideramos enriquecedora pois no mercado de trabalho problemas similares surgirão e temos de estar aptos a responder às divergências. A única parte que não correu tão bem foi o facto de o JDOM não ter funcionado.

7. Referências Bilbiográficas

- *github JDOM:* https://github.com/hunterhacker/jdom
- *jaxb:*https://github.com/javaee/jaxb-v2
- *jaxp :*https://www.oracle.com/java/technologies/jaxp-introduction.html
- *xml validator:*https://www.xmlvalidation.com/
- *xml validator w3:* https://validator.w3.org/
- *java xml tutorials:* *https://mkyong.com/tutorials/java-xml-tutorials/*
- *sax project:* http://www.saxproject.org/event.html

o teu • de partida

www.ipvc.pt