

Arquitecturas Orientadas a Servicios

Actividad 1: Procesando archivos XML

Problema:

Leer la información de un archivo XML mediante la librería DOM en Java, desplegar la información del archivo XML en un orden legible para el usuario mediante la consola.

Solución y explicación:

DOM es una librería para Java que permite extraer la información de un archivo XML mediante un objeto de la clase File de java.io.

Las clases de la librería DOM utilizadas en esta práctica fueron:

- DocumentBuilderFactory
- DocumentBuilder
- Document
- NodeList
- Node
- Element

Estas clases serán nombradas en el desarrollo del informe a medida que se vaya explicando el uso que se le dieron durante la práctica.

El objetivos de la práctica es obtener la información de un archivo XML del catastro para desplegarla en pantalla de manera legible para el usuario.

Para obtener la información del XML debemos usar la clase factoría de DOM (DocumentBuilderFactory) la cual está implementada bajo el patrón de diseño Factory y nos permite crear objetos de tipo DocumentBuilder a los cuales les pasaremos el objeto File para que esta pueda ser parseado a un objeto de tipo Document.

Una vez tenemos el Document necesitamos normalizar su contenido para obtener la información del archivo XML en un estándar, esto se hace para trabajar siempre sobre información normalizada, para de esta manera evitar espacios en blanco o demás inconsistencias que pueda tener el archivo.

Posterior a esto procedemos a extraer la información en estructuras de datos manipulables, es decir, que podamos iterar, consultar o en su defecto, operar.

Del Document podemos extraer un árbol con el contenido del archivo representado en listas de nodos de clase NodeList, cada nodo es un objeto de la clase Node y para poder extraer la información del nodo debemos manipularlo como objeto de la clase Element, para poder hacerlo, se hace un casting de Java de Node a Element.

Paso siguiente obtenemos el nodo raíz de nombre Provincias, para poder trabajar sobre los subnodos del árbol XML se hace el casting de Node a Element, de esta manera podemos descender en la jerarquía del árbol y podremos obtener la información de los atributos que deseamos.

Del nodo raíz obtenemos la dirección web sobre la cual se hizo la consulta soap y el número de consulta.

Seguimos avanzando en la jerarquía del archivo para poder obtener la información de las provincias del catastro.

Iteramos entonces sobre todos los nodos hijos del nodo provinciero, es decir, todos los nodos de *tagname* prov que estén por debajo del nodo provinciero.

Para verificar que estamos en el nodo correcto imprimimos el *tagname* sobre el cual estamos posicionados y seguido de eso imprimimos uno a uno la información obtenida de cada provincia. Esta información es: código postal y nombre.

Una vez ejecutada la aplicación la información en pantalla se muestra en el siguiente formato:

```
Element raíz:Provincias
Información obtenida de: http://www.catastro.meh.es/
Numero de Control de consulta: 48
```

```
-----
INFORMACIÓN CATASTRO
-----
```

```
Tipo de nodo actual :prov
Codigo Postal Provincia : 15
Nombre Provincia : A CORUÑA
```

De esta manera se despliegan todas las provincias obtenidas de la consulta soap.

Se adjunta una captura de pantalla evidenciando la ejecución de la aplicación:

```
1 import javax.xml.parsers.DocumentBuilderFactory;
2
3 public class XMLReader {
4
5     public static void main(String[] args){
6         try {
7             //Abrir el archivo xml
8             File XMLFile = new File("src/catastro2.xml");
9             //Obtenemos el xml en un Document DOM para poder procesarlo
10            Document DOM = DocumentBuilderFactory.newInstance().newDocumentBuilder().parse(XMLFile);
11            //Normalizamos la información para poder mostrarla en organizada
12            DOM.getDocumentElement().normalize();
13            //Imprimimos el elemento raíz del DOM
14            System.out.println("Element raíz:" + DOM.getDocumentElement().getNodeName());
15            //Obtenemos el nodo provincias para saber de donde obtenemos la información
16            Node nodeProvincias = DOM.getElementsByTagName("Provincias").item(0);
17
18        } catch (Exception e) {
19            e.printStackTrace();
20        }
21    }
22 }
```

<terminated> XMLReader [Java Application] /Library/Java/JavaVirtualMachines/jdk1.8.0_25.jdk/Contents/Home/bin/java (8/10/2015 7:33:19 p. m.)
Element raíz:Provincias
Información obtenida de: http://www.catastro.meh.es/
Numero de Control de consulta: 48

INFORMACIÓN CATASTRO

Tipo de nodo actual :prov
Codigo Postal Provincia : 15
Nombre Provincia : A CORUNA

Tipo de nodo actual :prov
Codigo Postal Provincia : 03
Nombre Provincia : ALACANT

Tipo de nodo actual :prov
Codigo Postal Provincia : 02
Nombre Provincia : ALBACETE

El código fuente puede encontrarse también en el directorio del curso en github:

<https://github.com/danielcb29/SOA>