**WHAT IS XML FILE?**

XML stands for Extensible Markup Language. It is a self-descriptive markup language like HTML which is basically used to store and transport data from one channel to another. XML is a W3C recommendation.

### ****POINTS TO PONDER****

* XML fully takes care of the data
* It is used to transport data
* There are no pre-defined tags in XML like HTML

### ****WHAT IS XML PARSER?****

An XML parser is a technique to access and manipulate XML files by loading its XML DOM.

### ****WHAT IS XML DOM?****

An XML DOM (Document Object Model) is a standard way to access and manipulate the message written inside XML documents. The entire DOM looks like the tree structure.

The info from XML DOM is retrieved by using JavaScript methods.

### ****WHAT IS XML NODE?****

Every element written inside an XML file is a node.

### ****WHAT IS NODE LIST?****

A list of nodes which consist of child node as well. Every element of the node is retrieved by using **getElementsByTagName()** method.

<https://www.inviul.com/parse-xml-file-selenium/>

### ****SAMPLE XML FILE****

Here is the sample XML file whose data we are going to read through Java for our Selenium project.

<?xml version = "1.0"?>

<bank>

<account acn = *"001"*>

<firstname>Avinash</firstname>

<lastname>kumar</lastname>

<balance>50000</balance>

</account>

<account acn = *"002"*>

<firstname>John</firstname>

<lastname>hari</lastname>

<balance>500</balance>

</account>

<account acn = *"003"*>

<firstname>Michel</firstname>

<lastname>dsouza</lastname>

<balance>1000</balance>

</account>

</bank>

In the above sample XML file, the bank is the head node and account is the child node. We will further get values of each node written inside the account node.

### ****TEST DATA PARSING FROM XML FILE IN SELENIUM WEBDRIVER****

**import** java.io.File;

**import** javax.xml.parsers.DocumentBuilder;

**import** javax.xml.parsers.DocumentBuilderFactory;

**import** org.w3c.dom.Document;

**import** org.w3c.dom.Element;

**import** org.w3c.dom.Node;

**import** org.w3c.dom.NodeList;

**public** **class** XMLFileReader {

**public** **static** **void** main(String[] args) {

**try**{

String filePath = "G:\\Selenium\\eclipse-workspace\\FirstDemo\\input.xml";

File file = **new** File(filePath);

//These two lines of code create the instance of the DOM architecture of the XML file.

DocumentBuilderFactory dbf = DocumentBuilderFactory.*newInstance*();

DocumentBuilder dBuilder = dbf.newDocumentBuilder();

//These 2 lines of codes parse the XML file and normalize it for test data retrieval.

Document doc = dBuilder.parse(file);

doc.getDocumentElement().normalize();

System.***out***.println(doc.getDocumentElement().getNodeName());

NodeList nodeList = doc.getElementsByTagName("account");

**int** tLength = nodeList.getLength();

**for**(**int** i=0; i<tLength; i++){

Node node = nodeList.item(i);

**if**(node.getNodeType()==Node.***ELEMENT\_NODE***){

Element element = (Element)node;

System.***out***.println("Account No: "+element.getAttribute("acn"));

System.***out***.println("First Name: "+element.getElementsByTagName("firstname").item(0).getTextContent());

System.***out***.println("Last Name: "+element.getElementsByTagName("lastname").item(0).getTextContent());

System.***out***.println("Balance: "+element.getElementsByTagName("balance").item(0).getTextContent());

}

}

}**catch** (Exception e){

e.printStackTrace();

}

}

}

**Output:**

bank

Account No: 001

First Name: Avinash

Last Name: kumar

Balance: 50000

Account No: 002

First Name: John

Last Name: hari

Balance: 500

Account No: 003

First Name: Michel

Last Name: dsouza

Balance: 1000

What is the use of getTextContent() and what is item(0)?

<https://stackoverflow.com/questions/22909160/what-does-gettextcontent-and-item0-return-in-java-dom-api>

By calling getTextContent() you are accessing inner text, by calling item(0) you are getting first accessible tag with such tagname.