

TIKA - EXTRACTING JAR FILE

http://www.tutorialspoint.com/tika/tika_extracting_jar_file.htm

Copyright © tutorialspoint.com

Given below is the program to extract content and metadata from a Java Archive *jar* file:

```
import java.io.File;
import java.io.FileInputStream;
import java.io.IOException;

import org.apache.tika.exception.TikaException;
import org.apache.tika.metadata.Metadata;
import org.apache.tika.parser.ParseContext;
import org.apache.tika.sax.BodyContentHandler;
import org.apache.tika.parser.pkg.PackageParser;

import org.xml.sax.SAXException;

public class PackageParse {

    public static void main(final String[] args) throws IOException, SAXException,
TikaException {

        //detecting the file type
        BodyContentHandler handler = new BodyContentHandler();
        Metadata metadata = new Metadata();
        FileInputStream inputstream = new FileInputStream(new File("Example.jar"));
        ParseContext pcontext = new ParseContext();

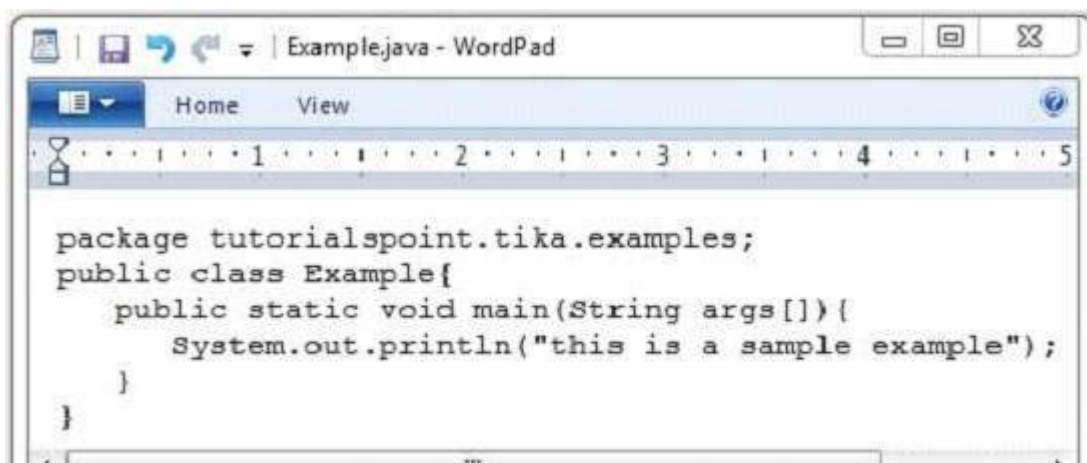
        //Package parser
        PackageParser packageparser = new PackageParser();
        packageparser.parse(inputstream, handler, metadata, pcontext);
        System.out.println("Contents of the document: " + handler.toString());
        System.out.println("Metadata of the document:");
        String[] metadataNames = metadata.names();

        for(String name : metadataNames) {
            System.out.println(name + ": " + metadata.get(name));
        }
    }
}
```

Save the above code as **PackageParse.java**, and compile it from the command prompt by using the following commands:

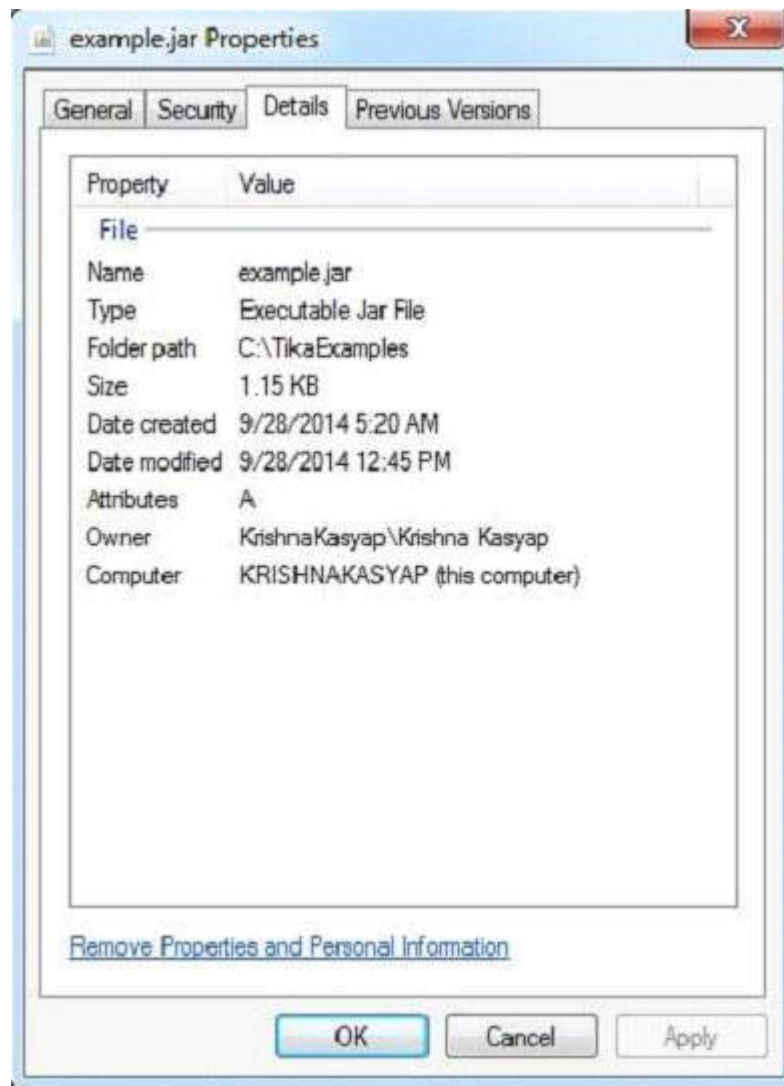
```
javac PackageParse.java
java PackageParse
```

Here we are passing the jar file of the following Example.java file.





The jar file has the following properties:



After executing the above program you will get the following output.

Output:

```
Contents of the document:
META-INF/MANIFEST.MF
tutorialspoint/tika/examples/Example.class
```

```
Metadata of the document:
```

```
Content-Type: application/zip
```

```
Loading [Mathjax]/jax/output/HTML-CSS/jax.js
```