

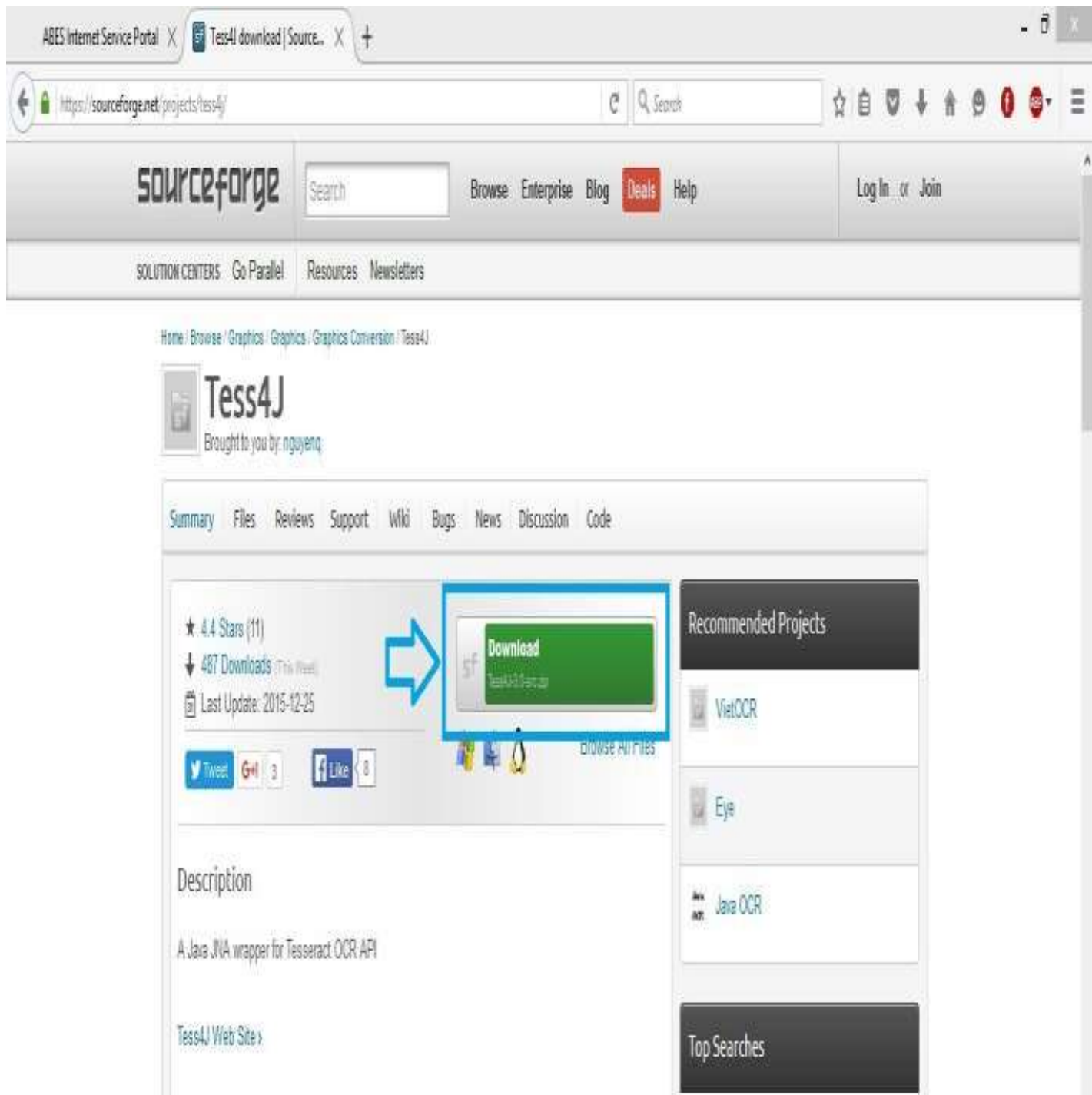
Installation of Tesseract-OCR in java on Window

Prerequisites:

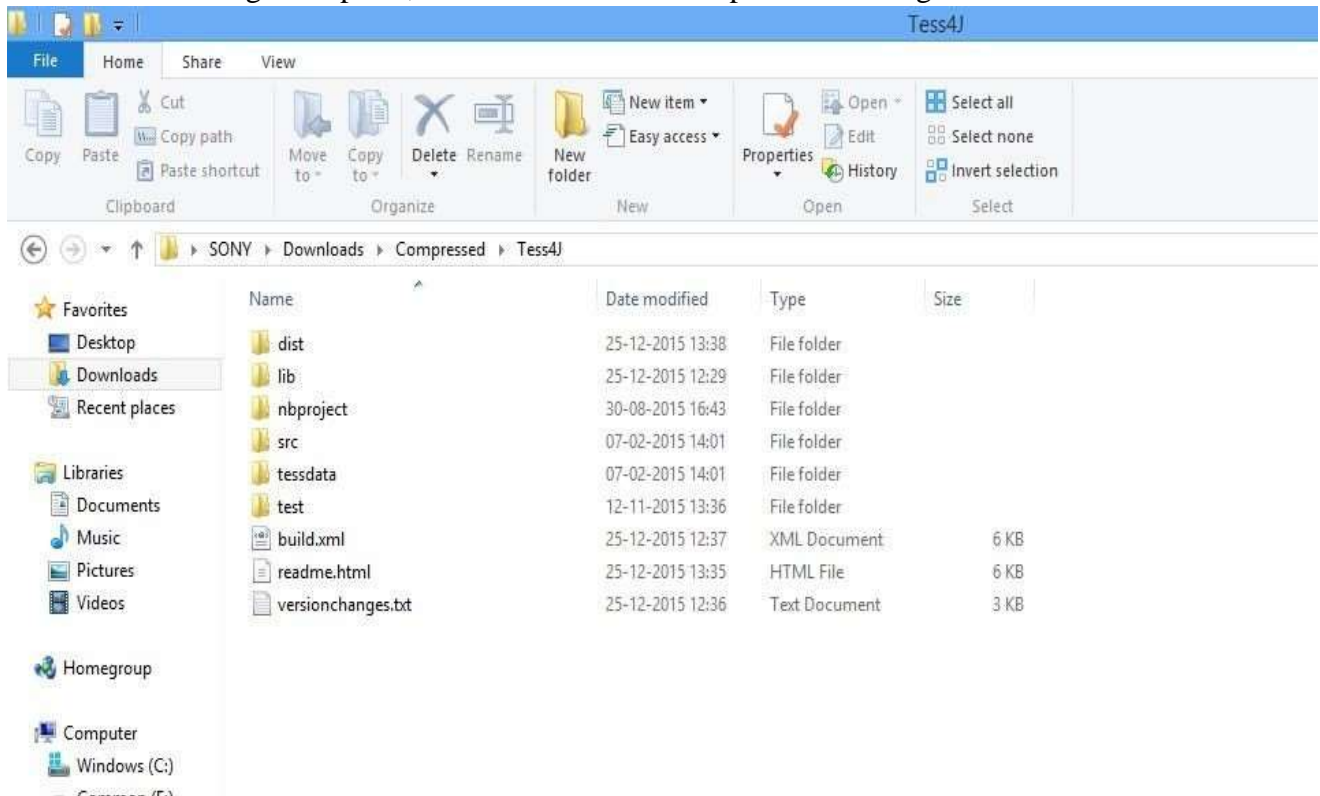
I am expecting that you have already installed Java and Eclipse .

Steps to install Tesseract in java on Window

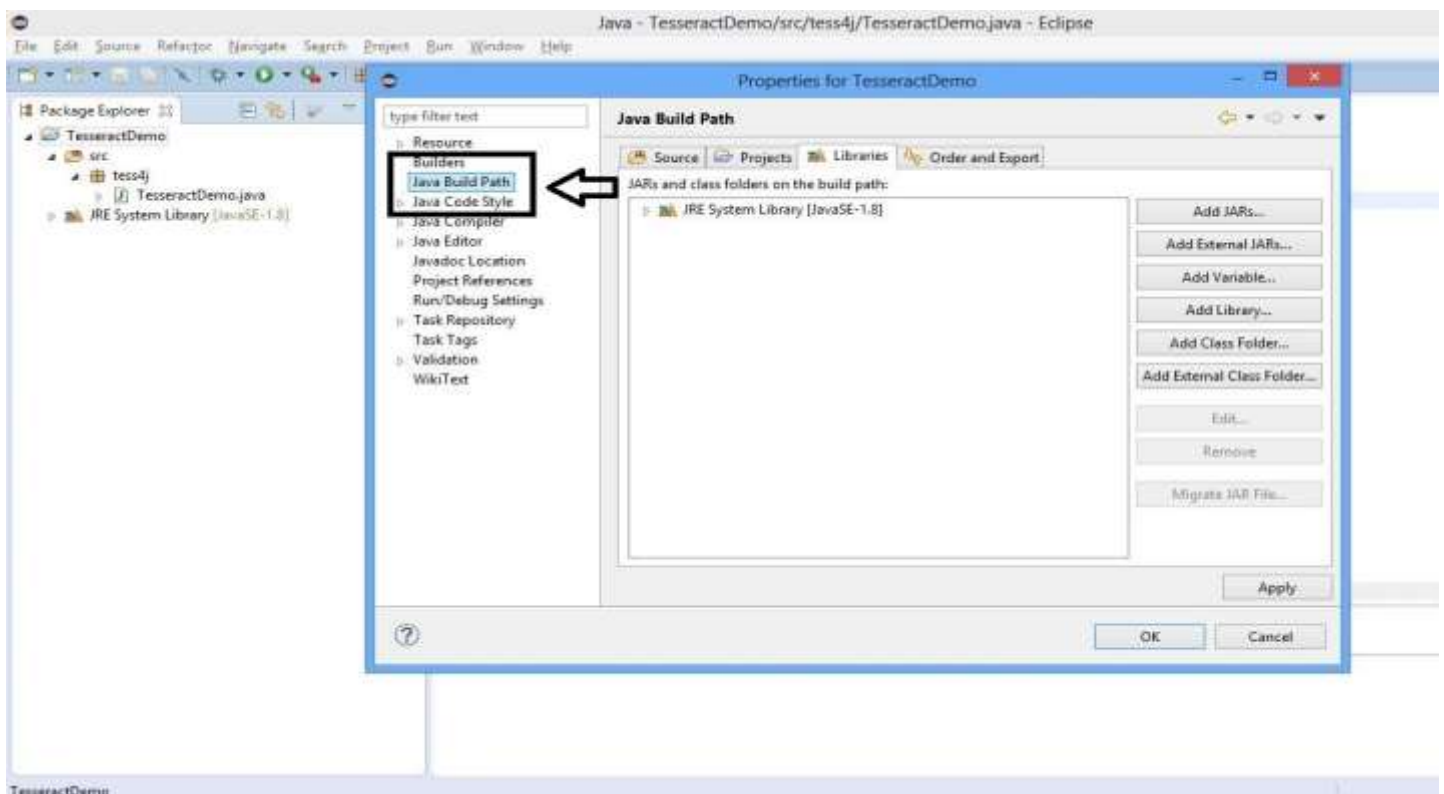
- Open this link <https://sourceforge.net/projects/tess4j/> to download tesseract
- Click on download button. I am using **IDM** to download it.



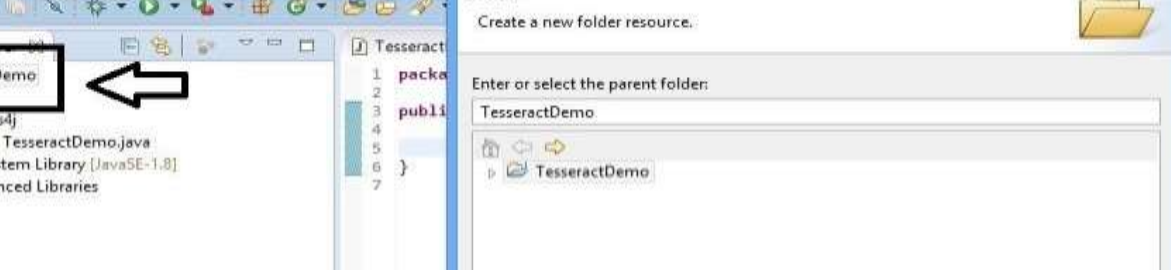
- After downloading the zip file, Extract Tess4J-3.0-src.zip and following folder will be look like this.



- Open eclipse and create a new project
- Right click on **project** (TesseractDemo) and then Click on Java Build Path
- Click on Configure Build Path
- In java Build path, Click on **libraries** and then click on Add External Jar
- Find and add all the required JAR files, including jai_imageio.jar, jna.jar, commons-io-2.4.jar, and tess4j.jar. These external Jar file is in **dist** and **lib** folder in tesseract Add a new *Java Class* file named TesseractDemo with appropriate package name

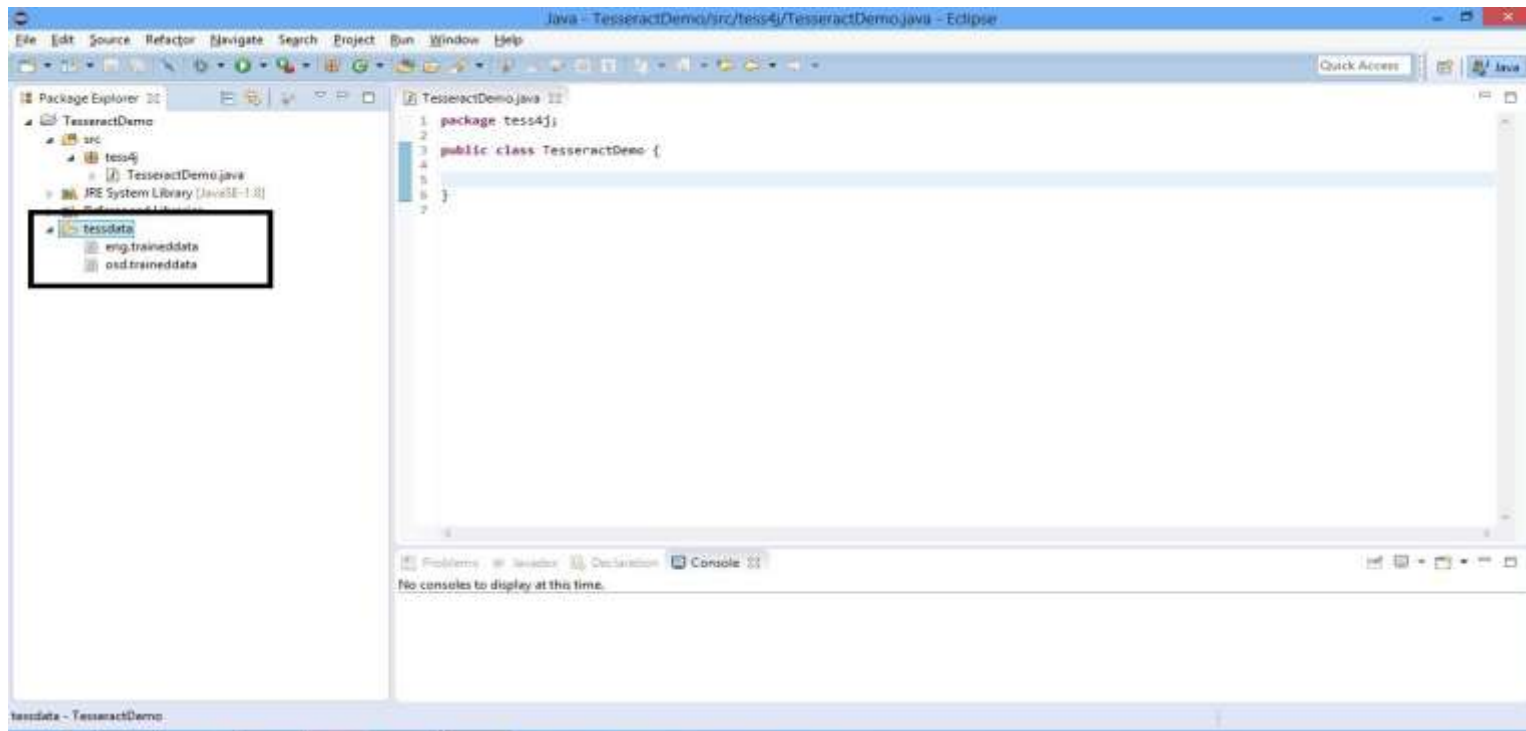


- [illegible]

- 
- The screenshot shows an IDE with the 'New Folder' dialog open. The dialog has a title bar 'New Folder' and a subtitle 'Create a new folder resource.' Below this, there is a text field 'Enter or select the parent folder:' with 'TesseractDemo' entered. A list of folders is shown below, with 'TesseractDemo' selected. At the bottom, there is a text field 'Folder name:' with 'tessdata' entered. There are 'Advanced >>' and 'Finish' buttons. In the background, the IDE's project structure is visible, showing 'TesseractDemo' with a 'src' folder and 'TesseractDemo.java' file. A red arrow points to the 'TesseractDemo' folder in the project structure.

Note:

Now **copy eng.traineddata and osd.trained data** from **tess4j/tessdata** to your **Eclipse_workspace/your_project(say tesseractDemo)/tessdata**



Testing the configuration

```
package tess4j;
import java.io.File;
import net.sourceforge.tess4j.*;
public class TesseractDemo {
public static void main(String[] args) {
    System.setProperty("jna.library.path", "32".equals(System.getProperty("sun.arch.data.model")) ? "lib/win32-x86" : "lib/win32-x86-64");
    File imageFile = new File("eurotext.tif");
    ITesseract instance = new Tesseract(); // JNA Interface Mapping
    // ITesseract instance = new Tesseract1(); // JNA Direct Mapping
    // File tessDataFolder = LoadLibs.extractTessResources("tessdata"); // Maven build bundles English data
    // instance.setDatapath(tessDataFolder.getAbsolutePath());
    try {
        String result = instance.doOCR(imageFile);
        System.out.println(result);
    } catch (TesseractException e) {
        System.err.println(e.getMessage());
    }
}
```

After running the project, result is presented on console

