

What are Broken Links?

Broken links are links or URLs that are not reachable. They may be down or not functioning due to some server error

An URL will always have a status with 2xx which is valid. There are different HTTP status codes which are having different purposes. For an invalid request, HTTP status is 4xx and 5xx.

4xx class of status code is mainly for client side error, and 5xx class of status codes is mainly for the server response error.

Why should you check Broken links?

You should always make sure that there are no broken links on the site because the user should not land into an error page.

The error happens if the rules are not updated correctly, or the requested resources are not existing at the server.

Manual checking of links is a tedious task, because each webpage may have a large number of links & manual process has to be repeated for all pages.

How to check the Broken Links and Images in Selenium

For checking the broken links, you will need to do the following steps.

1. Collect all the links in the web page based on <a> tag.
2. Send HTTP request for the link and read HTTP response code.
3. Find out whether the link is valid or broken based on HTTP response code.
4. Repeat this for all the links captured.

```

package practice;
import java.net.HttpURLConnection;
import java.net.MalformedURLException;
import java.net.URL;
import java.util.List;
import org.openqa.selenium.By;
import org.openqa.selenium.WebDriver;
import org.openqa.selenium.WebElement;
import org.openqa.selenium.chrome.ChromeDriver;
public class brokenlink {
    public static void main(String[] args) {
        WebDriver driver = new ChromeDriver();
        driver.manage().window().maximize();
        driver.get("http://www.deadlinkcity.com/");

        List<WebElement> linkList = driver.findElements(By.tagName("a")); //Collect
all the links in the web page based on <a> tag.
        System.out.println("Total links on page =" + linkList.size());
        int resCode = 200; //2xx Responce code is valid
        int brokenLinkCnt = 0;

        for(WebElement element : linkList)
        {
            String url = element.getAttribute("href");

            try {
                URL urlLink = new URL(url);
                HttpURLConnection huc =
(HttpURLConnection)urlLink.openConnection();
                huc.setRequestMethod("HEAD");
                huc.connect();
                resCode = huc.getResponseCode();

                if(resCode >= 400) //4xx and 5xx series are broken link
                {
                    System.out.println(url + " : broken Link" );
                    brokenLinkCnt++;
                }
            }
            catch(MalformedURLException e)
            {
            }
            catch(Exception e)
            {
            }
        }
        System.out.println("Total Broken Link =" + brokenLinkCnt );
    }
}

```

eclipse-workspace_2 - Interview/src/main/java/practice/brokenlink.java - Eclipse IDE
File Edit Source Refactor Navigate Search Project Run Window Help

Console ×

```
<terminated> brokenlink [Java Application] C:\Users\Amol\p2\pool\plugins\org.eclipse.justi.openjdk.hotspot.jre.full.win32.x86_64.17.0.4.v20221004-1257\jre\bin\javaw.exe (20-Aug-2023, 10:55:40 am - 1
http://www.deadlinkcity.com/error-page.asp?e=429 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=431 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=450 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=500 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=501 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=502 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=503 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=504 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=505 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=506 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=507 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=509 : broken Link
http://www.deadlinkcity.com/error-page.asp?e=510 : broken Link
http://www.deadlinkcity.com/page-not-found.asp : broken Link
http://www.deadlinkcity.com/default.asp?r=4 : broken Link
http://www.deadlinkcity.com/default.asp?r=5 : broken Link
http://www.deadlinkcity.com/default.asp?r=6 : broken Link
http://www.deadlinkcity.com/default.asp?r=7 : broken Link
http://www.deadlinkcity.com/disallowed/disallowed.html : broken Link
http://www.deadlinkcity.com/disallowed/nonexistent.html : broken Link
Total Broken Link =42
```