

README:

HTML Parser Used: JTidy

JDK Version: 1.7.0

Compile and Run Indexer:

1. run javac and specify external jars

```
javac -cp "./Lucene/*:./HTMLParser/*" Indexer.java
./HTMLParser/JTidyHTMLHandler.java
```

2. run program

```
java -cp ./Lucene/*:./HTMLParser Indexer -index ./_index -docs
./_webpagesToIndex/
```

Compile and Run Searcher:

1. run javac and specify external jars

```
javac -cp "./Lucene/*" Searcher.java
```

2. run program

```
java -cp ./Lucene/* Searcher -index ./_index -q "machine learning"
```

Reset (delete .class files and all output for indexer/searcher):

```
rm *.class; rm ./HTMLParser/JTidyHTMLHandler.class; rm -rf _index; rm
results.html
```