bin/nutch indexchecker

indexchecker is an alias for running the class org.apache.nutch.indexer.IndexingFiltersChecker

This tool fetches and parses an URL and then runs the indexing filters on it. Once complete, it displays the fields obtained and the first 100 characters of their value.

Usage:

```
bin/nutch IndexingFiltersChecker <url>
```

<url>: The URL you wish to run the indexers on.

The parameter *-D doIndex=true* can be specified either on the command line or in nutch-site.xml in order to send the document to the indexing backends. Those must be configured accordingly as described in the documentation for the index command. **It should be noted however that the <url> needs to be the final argument passed to the command line**

CommandLineOptions