

bin/nutch readlinkdb

Readlinkdb is an alias for `org.apache.nutch.crawl.LinkDbReader`

This reader class enables us to obtain various information from within a linkdb. The two types of information we can retrieve is

- A dump of the whole linkdb which is then written to a text file for easy viewing.
- Specific information relating to a specific URL.

[blocked URL](#) :TODO: More could be added to the above e.g what is the nature and structure of the information we retrieve from a dump of the linkdb and a specific URL. [blocked URL](#)

Usage:

```
bin/nutch readlinkdb <linkdb> (-dump <out_dir> | -url <url>)
```

<linkdb>: This is the linkdb directory we wish to read and obtain information from.

-dump <out_dir>: This parameter dumps the whole linkdb to a text file in any <out_dir> we wish to specify.

-url <url>: The -url argument provides us with information about a specific <url>. This is written to System.out.

[CommandLineOptions](#)