

Read me

Word Count Service

Contents

- Introduction
- How to build and run
- Performance improvement included in the service

Introduction

This is a JAVA service build on Grizzly and Jersey framework. It runs as a service and returns frequency of a given word present in a corpus in json format. Here corpus means all the text files present inside a folder.

How to build and run

- i. Import the project in the eclipse.
- ii. Right click on **build.xml** file present inside the **data** folder under the root folder and run as an ant application.
- iii. Above step creates a **WordCountService.zip** file inside the bin folder under the root folder.

For running, first unzip the zipped folder. Set following configuration parameters present in the **config.properties** file inside the **config** folder under root folder:

- **hostAddress=10.192.29.41:8089** (<Address/Hostname of machine>:<Port>)
- **crossdomainFilepath=crossdomain.xml**
- **corpusPath=data** (Folder where text files are present)

Then run the WordCountService.jar using following command:

```
java -jar WordCountService.jar
```

For testing, following get request:

<http://<IP/Hostname>:<Port>/getCount?query=<word>>

It returns count of the <word> present in all the files inside data folder in following json format:

```
{ "count" : "108" }
```

Performance improvement included in the service

This application caches the count of a word in a given file so that if the request for the same word is made for the 2nd time, there will be significant improvement in the response time. It also handles the case in which a file has been modified or a file has been added or deleted.