

Documentation for download tools at CRCNS.org.

Version 0.5 (Oct. 28, 2014)

This directory contains scripts to batch download CRCNS.org data files that are located at the National Energy Research Scientific Computing Center (NERSC) and to also verify the downloaded files.

There are two scripts:

<i>download.sh</i>	- downloads data set files
<i>verify.sh</i>	- verifies downloaded files

Both scripts run using bash (executed on a linux or mac command line).

To use the download and verify scripts:

Step 1.

Edit file: "crcns-account.txt" to specify an account (username and password) at CRCNS.org.

Step 2.

Change permission on the scripts to allow executing. This is done using the command:

```
chmod u+x download.sh verify.sh
```

Step 3.

Run the download script by typing:

```
./download.sh <dataset_id>
```

Where <dataset_id> is the id of the data set, for example, "pvc-1" (without the quotes).

If the dataset_id is not specified, the script will prompt for it.

The script will create a directory with the name of the dataset id, and download a file named "filelist.txt" into the directory.

Step 4.

Edit the file "filelist.txt" by following the instructions at the top to specify which files in the dataset to download.

This will require putting either a "#" or "+" character in the first position of lines to specify files to exclude (default mode) or to download (+ mode).

Step 5.

Rerun the script as in step 3 to download the files.

To download additional files, edit the filelist.txt file and rerun the script.

The script and server support resume downloads, so that if a download does not complete, just rerun the script again and the download will resume.

Step 6.

To verify the files were downloaded correctly, download file "checksums.md5" for the data set (using the download script), then run the verify script, by typing:

```
./verify.sh <dataset_id>
```

This will verify the downloaded files using md5 checksums.

To provide feedback or about these tools, send email to "admin@crcns.org".