Dammmien / **wget.sh**

Last active 2 months ago • Report abuse

<> **Code**   -o- Revisions 2   ☆ Stars 53   🖇 Forks 15

---

wget cheat sheet

<> **wget.sh**

```
1    # POST a JSON file and redirect output to stdout
2    wget -q -O - --header="Content-Type:application/json" --post-file=foo.json http://127.0
3
4    # Download a complete website
5    wget -m -r -linf -k -p -q -E -e robots=off http://127.0.0.1
6
7    # But it may be sufficient
8    wget -mpk http://127.0.0.1
9
10   # Download all images of a website
11   wget -r -l inf -A .png,.jpg,.jpeg,.gif -nd https://jekyllrb.com
12
13   # FTP download
14   wget --ftp-user=User --ftp-password=Mir URL
15
16
17   Startup:
18     -V,  --version           display the version of Wget and exit.
19     -h,  --help              print this help.
20     -b,  --background        go to background after startup.
21     -e,  --execute=COMMAND   execute a `.wgetrc'-style command.
22
23   Logging and input file:
24     -o,  --output-file=FILE    log messages to FILE.
25     -a,  --append-output=FILE  append messages to FILE.
26     -d,  --debug               print lots of debugging information.
27     -q,  --quiet               quiet (no output).
28     -v,  --verbose             be verbose (this is the default).
29     -nv, --no-verbose          turn off verboseness, without being quiet.
30          --report-speed=TYPE   Output bandwidth as TYPE.  TYPE can be bits.
31     -i,  --input-file=FILE     download URLs found in local or external FILE.
32     -F,  --force-html          treat input file as HTML.
33     -B,  --base=URL            resolves HTML input-file links (-i -F)
34                                relative to URL.
35          --config=FILE         Specify config file to use.
36
37   Download:
38     -t,  --tries=NUMBER         set number of retries to NUMBER (0 unlimits).
39          --retry-connrefused    retry even if connection is refused.
40     -O,  --output-document=FILE  write documents to FILE.
```

```
41      -nc, --no-clobber                 skip downloads that would download to
42                                         existing files (overwriting them).
43      -c,  --continue                   resume getting a partially-downloaded file.
44           --progress=TYPE              select progress gauge type.
45      -N,  --timestamping               don't re-retrieve files unless newer than
46                                         local.
47      --no-use-server-timestamps        don't set the local file's timestamp by
48                                         the one on the server.
49      -S,  --server-response            print server response.
50           --spider                     don't download anything.
51      -T,  --timeout=SECONDS            set all timeout values to SECONDS.
52           --dns-timeout=SECS           set the DNS lookup timeout to SECS.
53           --connect-timeout=SECS       set the connect timeout to SECS.
54           --read-timeout=SECS          set the read timeout to SECS.
55      -w,  --wait=SECONDS               wait SECONDS between retrievals.
56           --waitretry=SECONDS          wait 1..SECONDS between retries of a retrieval.
57           --random-wait                wait from 0.5*WAIT...1.5*WAIT secs between retrievals.
58           --no-proxy                   explicitly turn off proxy.
59      -Q,  --quota=NUMBER               set retrieval quota to NUMBER.
60           --bind-address=ADDRESS       bind to ADDRESS (hostname or IP) on local host.
61           --limit-rate=RATE            limit download rate to RATE.
62           --no-dns-cache               disable caching DNS lookups.
63           --restrict-file-names=OS     restrict chars in file names to ones OS allows.
64           --ignore-case                ignore case when matching files/directories.
65      -4,  --inet4-only                 connect only to IPv4 addresses.
66      -6,  --inet6-only                 connect only to IPv6 addresses.
67           --prefer-family=FAMILY       connect first to addresses of specified family,
68                                         one of IPv6, IPv4, or none.
69           --user=USER                  set both ftp and http user to USER.
70           --password=PASS              set both ftp and http password to PASS.
71           --ask-password               prompt for passwords.
72           --no-iri                     turn off IRI support.
73           --local-encoding=ENC         use ENC as the local encoding for IRIs.
74           --remote-encoding=ENC        use ENC as the default remote encoding.
75           --unlink                     remove file before clobber.
76
77   Directories:
78     -nd, --no-directories            don't create directories.
79     -x,  --force-directories         force creation of directories.
80     -nH, --no-host-directories       don't create host directories.
81          --protocol-directories      use protocol name in directories.
82     -P,  --directory-prefix=PREFIX   save files to PREFIX/...
83          --cut-dirs=NUMBER           ignore NUMBER remote directory components.
84
85   HTTP options:
86          --http-user=USER          set http user to USER.
87          --http-password=PASS      set http password to PASS.
88          --no-cache                disallow server-cached data.
89          --default-page=NAME       Change the default page name (normally
90                                     this is `index.html'.).
91     -E,  --adjust-extension        save HTML/CSS documents with proper extensions.
92          --ignore-length           ignore `Content-Length' header field.
```

```
 93           --header=STRING         insert STRING among the headers.
 94           --max-redirect          maximum redirections allowed per page.
 95           --proxy-user=USER       set USER as proxy username.
 96           --proxy-password=PASS   set PASS as proxy password.
 97           --referer=URL           include `Referer: URL' header in HTTP request.
 98           --save-headers          save the HTTP headers to file.
 99      -U,  --user-agent=AGENT      identify as AGENT instead of Wget/VERSION.
100           --no-http-keep-alive    disable HTTP keep-alive (persistent connections).
101           --no-cookies            don't use cookies.
102           --load-cookies=FILE     load cookies from FILE before session.
103           --save-cookies=FILE     save cookies to FILE after session.
104           --keep-session-cookies  load and save session (non-permanent) cookies.
105           --post-data=STRING      use the POST method; send STRING as the data.
106           --post-file=FILE        use the POST method; send contents of FILE.
107           --content-disposition   honor the Content-Disposition header when
108                                   choosing local file names (EXPERIMENTAL).
109           --content-on-error      output the received content on server errors.
110           --auth-no-challenge     send Basic HTTP authentication information
111                                   without first waiting for the server's
112                                   challenge.
113
114   HTTPS (SSL/TLS) options:
115           --secure-protocol=PR    choose secure protocol, one of auto, SSLv2,
116                                   SSLv3, and TLSv1.
117           --no-check-certificate  don't validate the server's certificate.
118           --certificate=FILE      client certificate file.
119           --certificate-type=TYPE client certificate type, PEM or DER.
120           --private-key=FILE      private key file.
121           --private-key-type=TYPE private key type, PEM or DER.
122           --ca-certificate=FILE   file with the bundle of CA's.
123           --ca-directory=DIR      directory where hash list of CA's is stored.
124           --random-file=FILE      file with random data for seeding the SSL PRNG.
125           --egd-file=FILE         file naming the EGD socket with random data.
126
127   FTP options:
128           --ftp-user=USER         set ftp user to USER.
129           --ftp-password=PASS     set ftp password to PASS.
130           --no-remove-listing     don't remove `.listing' files.
131           --no-glob               turn off FTP file name globbing.
132           --no-passive-ftp        disable the "passive" transfer mode.
133           --preserve-permissions  preserve remote file permissions.
134           --retr-symlinks         when recursing, get linked-to files (not dir).
135
136   WARC options:
137           --warc-file=FILENAME     save request/response data to a .warc.gz file.
138           --warc-header=STRING     insert STRING into the warcinfo record.
139           --warc-max-size=NUMBER   set maximum size of WARC files to NUMBER.
140           --warc-cdx               write CDX index files.
141           --warc-dedup=FILENAME    do not store records listed in this CDX file.
142           --no-warc-compression    do not compress WARC files with GZIP.
143           --no-warc-digests        do not calculate SHA1 digests.
144           --no-warc-keep-log       do not store the log file in a WARC record.
```

```
145              --warc-tempdir=DIRECTORY  location for temporary files created by the
146                                        WARC writer.
147
148    Recursive download:
149      -r,  --recursive          specify recursive download.
150      -l,  --level=NUMBER        maximum recursion depth (inf or 0 for infinite).
151           --delete-after        delete files locally after downloading them.
152      -k,  --convert-links       make links in downloaded HTML or CSS point to
153                                     local files.
154      -K,  --backup-converted    before converting file X, back up as X.orig.
155      -m,  --mirror              shortcut for -N -r -l inf --no-remove-listing.
156      -p,  --page-requisites     get all images, etc. needed to display HTML page.
157           --strict-comments     turn on strict (SGML) handling of HTML comments.
158
159    Recursive accept/reject:
160      -A,  --accept=LIST              comma-separated list of accepted extensions.
161      -R,  --reject=LIST              comma-separated list of rejected extensions.
162           --accept-regex=REGEX       regex matching accepted URLs.
163           --reject-regex=REGEX       regex matching rejected URLs.
164           --regex-type=TYPE          regex type (posix).
165      -D,  --domains=LIST             comma-separated list of accepted domains.
166           --exclude-domains=LIST     comma-separated list of rejected domains.
167           --follow-ftp               follow FTP links from HTML documents.
168           --follow-tags=LIST         comma-separated list of followed HTML tags.
169           --ignore-tags=LIST         comma-separated list of ignored HTML tags.
170      -H,  --span-hosts              go to foreign hosts when recursive.
171      -L,  --relative               follow relative links only.
172      -I,  --include-directories=LIST  list of allowed directories.
173      --trust-server-names              use the name specified by the redirection
174                                        url last component.
175      -X,  --exclude-directories=LIST  list of excluded directories.
```