



wget - руководство GNU Wget

GNU Wget

????????

wget -- ????????????? GNU Wget

????????

wget [????????]... [URL]...

????????

GNU Wget ??? ?????? ?????????????????? ??????? ?? ????????? ?????? ?? ?????????.
??? ??????????????
?????????? **HTTP**, **HTTPS**, ? **FTP**, ????????? ? ???????-????????? ?? ?????????? **HTTP**.

Wget ?????? ?????????? ?? ????????? HTML ????????? ? ?????????? ?????????? ?????? ??????
????? web ??????,
????????? ?????? ?????????????????? ?????????? ?????????? ??????. ??? ?????????????? "
recursive downloading"
(????????????? ??????????). ??? ?????????????? ????????? ?????????, Wget ?????? ?????? ? ?????????
? ?????????? ???
????????? (/robots.txt
) . ?????????? ?????? ?????????????? ????????? ? ?????????????? HTML ?????????? ??? ??????????
?????? ? ?????????????? ????????? ("Off-line browsing").

????????? ?????????????? ?????????? . Wget ?????? ?????????????? ?????????????? ??????, ??? ??????????
?? ??????????????

HTTP ? FTP

? ?????????? ? ?????????????? ??????, ?????????? ??? ?????????? ??????, ?, ?????? ??? ???, ?
?????????
?????? ?????????? ??????.
????????????? ??????, ? Wget ?????????? ?????????????? ?????????????? ?????????? ? ?????????????? ?????????
?? **FTP**.

Wget ?????????????? ??? ?????????????? ??? ?????????????????? ??????????????: ?????? ?? ?????? ??????
??? ??????????????
????????????, Wget ?????? ?????????????? ?????????????? ?????????????? ??????. ?????? ?????????, ? ?????????
?? ??????????????????
?????, ?????????????????? ??????????, Wget ?????????????? ?????????????? ?????? ? ?????? ??????, ??? ??
????????????? ?????????????.

O?????



????????? ??????????

-V

--version

????????????? ?????? Wget.

-h

--help

????????????? ?????????? ?????????? ?????? Wget.

-b

--background

????????? ? ?????????? ?????? ?????? ??????????. ????? ????? ??? ?????????? ?? ??????? ??????
?????? **-o**, ??
????????????????? ? **wget-log**

-e command

--execute command

???????????? **command** ??? ????? ?? ??? ????? ?????? **.wgetrc**
. ????????? ?????? ????????????? ?????? ?????????
? **.wgetrc**.

????????????? ?????????????

-o logfile

--output-file=logfile

????????????????? ??? ?????????????? ? **logfile**. ?????? ??? ?????? ?????????????? ? **stderr**.

-a logfile

--append-output=logfile

????????????? **logfile**. ??? ? **-o**, ??????? **logfile**
?? ??????????????, ? ??????????????. ???? **logfile**
?? ??????????????, ????????????? ?????? ??????

-d

--debug

????????????????? ?????????????? ?????????? - ?????????? ?????????????????, ?????????? ??? ?????????????????? Wge
t.

-q

--quiet

????????????? ?????????????? Wget.

-v

--verbose

????????????? ?????????????? ??????????????, ?? ??????? ?????????????????? ??????????. ?? ?????????????? ?????????
??.

-nv

--non-verbose

?? ?????????????? ?????????????? (?????? ?????????????? ?????????????? ??). **-q**
) . ?????????????? ?? ?????????? ? ??????????
????????????????? ?????? ??????????????????

-i file

--input-file=file

????????????? URL ?? file. ? ?????? ?????????? ?????????????? URL ? ?????????????? ?????????? ?? ??????
?????. ???? URL
????????????? ? ? ?????????????? ?????????? ? ? file, ?????????? ?????????? ?????????????? URL ?? ??????????????



? ??????. file
?? ?????????????? ?????? ?????? ?????? HTML (?? ??????? ????????????, ????? ??? ???) --
URL ???????
?????? ?????? ?????????? ? ???.

????? ?? ?????????? **--force-html**
, ????? ?????? ?????????? ??? html. ? ????? ?????????? ?????? ??????????????
?????????? ? ?????????????????????? ???????????. ??? ?????? ?????????????????????? ?????????????????? ? ??????

"<base href="url">" ??? ??????? ? ?????????????? ?????????? **--base=url**.

-F

--force-html

??? ??????? URL ?? ??????, ?????????? ??????? ?????? ??? HTML. ?????? ?????????????????? ?
????????, ????
????? HTML ??????????? ??????????????, ??????????? "<base href="url">
" ? ????? ??? ?????????? ??????????
????????????? ?????????? **--base**.

-B URL

--base=URL

??? ??????? URL ?? ?????? (-F
) , ?????????????? URL, ?????????????????? ? ?????????????????????? ?????????? ??????,
????????????????? ?????????????????? **-i**.

????????????? ??????????????

--bind-address=ADDRESS

??? ?????????????????? ?? ?????????????????? TCP/IP, ??????????? "bind()" ? ADDRESS
?? ?????????????????? ??????????
? ADDRESS
?????? ?????? ?????????? ??? ??? ??????, ??? ? IP ??????. ??????????????????, ?????? ???
????????????????? ?????? ?????????????????? IP ??????????.

-t number

--tries=number

????????????????????? ?????? ?????????????? number. ?????????? 0 ??? inf
??? ?????????????????????? ?????? ??????????????

-O file

--output-document=file

?????????????? ?? ?????? ?????????????????????? ? ?????????????????????????????? ??????, ? ?????? ?????????????????? ???
??? ?
????????????? ? file. ?????? file ??????????????????, ?? ?????? ??????????. ?????? file
????????? ??? -,
????????????? ?????? ?????????????????? ? ?????????????????????? ?????? (stdout
). ?????? ?????????????? ??????????????????????
????????????????????? ?????? ?????????????? ? 1. ?????????? ??? ?????????????? ?????????, ?????????????????????? ?? ?
?????,
? ?????????????? ?????????????? ?????? web-?????????????????

-nc

--no-clobber

????? ??? ?????????????? ?????? ?????? ??????????????????, ??? ?????????????????????? ? ?????? ?? ??????, ???
?????
????? ??????????????

??? ?????????? Wget ??? **-N, -nc**
, ??? -r, ?????????????? ?????????????? ?????????? ? ?????? ?? ?????? ? ??????



???????? ???? ????? ? ????????? ????? ????? ? ?????? file.1. ??? ?????????? ???
 ? ? ?????
 ??????, ????? ????? ????? ????????????? file.2 ? ?. ?. ??? ?????????? **-nc**
 ??? ?????
 ????????????????

??? ???????? Wget ? ????????????? **-r**, ?? ??? **-N** ??? **-nc**
 , ????? ????????? ????? ????????? ?
 ?????? ??? ????????????? ??????. ??? ????????????? **-nc**
 ????????? ????????????? ? ?????? ??????
 ? ????????????? ?????? ?? ?????? ????????????? ?????? (???? ?????? ??? ? ? ????????????)
).

??? ???????? Wget ? ????????????? **-N**
 , ? ??? ??? **-r**, ????? ?????? ?????????? ????? ? ? ??????,
 ??? ????? ?????? ? ? ?????????? ? ????????????? ?????? (?. ?????????? ? ? ?????).
-nc ??
 ????????????????? ? **-N**.

??? ????????????? ????????????? **-nc**
 , ????? ? ????????????????? .html ??? (???? ?????? ??????) .htm
 ? ?????????? ?????? ?????? ????????????????? ? ? ? ??????????.

-c
--continue
 ????????????????? ?????????? ??????. ????????????????? ????? ?????????? ?????? ?????? ??????????
 ??????????:

wget -c <ftp://sunsite.doc.ic.ac.uk/ls-lR.Z> [1]

???? ? ???????? ?????????? ????? ????? ? ?????? **ls-lR.Z**
 , Wget ?????????, ????????????????? ?
 ?????? ????? ????????????????? (?? ?? ?????????!), ?, ????? ??? ???, ?????????? ??????? ? ?
 ??????
 ?? ?????????????? ?????????? ?????? ? ????? ? ? ?????, ??? ?????????????? ?????????? ? ??????
 ?? ???.

????????, ??? ??? ?????? ?????? Wget ?????????? ?????????? ?????????????????? ? ??? **-c**,
 ? ??????
 ?????? ?? "????????" ? ?????????? ??????, ??? ?????????????????? ?????? ?????? ?????? ??????
 ???.

??? ?????? **-c**
 , ????????????? ??????? ?????????? ? ?????????? ?????????????? ?????? ?????? ? ??????????
 ?????? **ls-lR.Z.1**, ?? ?????? **ls-lR.Z**

???????? ? ?????? 1.7 ??? ????????????? **-c**
 , ??? ???? ? ? ?????????? ?????? ??????, ??? ?????????,
 ??? ? ?????????????? ?????? ??????, Wget ?? ?????? ?????? ?????????? ? ?????????? ??????
 ??????????????
 ????????????

??????? ??? ?????????????????? **-c**
 , ?????? ???? ? ? ?????????, ?????????? ?????????? ??????, ??? ???????????,
 ?????? ?????????????????????, ??? ??????????????????. ??? ???? ?????? ????????????? ? ?????????? ?
 ?????? ??????
 ?????? "(?????(?????.????) - ??????(?????.????))" ??????. ??? ?????? ??????????????, ?
 ?????? ???



????? ?????????? ?????? ?????????? ??????-????? ???????? (log).

??? ?????, ????? ?????????? ????? ??????? ???????, ??? ?? ??????????, ?? ?????????? ???
????????????

???? (?.?. ????? ????????????? ??????). ??? ?????? ???? ?????????? ?????????????????? ???
????????????????? -c ?????? ? -r, ??? ??? ??????? ?????????????? ????? ?????? ????? ??????
????? ??
"????????????????? ??????????".

?? ?????? ?????????? ?????????????????? ?????, ????? ??? HTTP ??????-???????? ?????????????? ?
?????????
? ??? ??????? ?????????????? ?????? "**transfer interrupted**
" ? ??????. ? ????????????? ?????????? Wget
??? ?????? ?????????????? ???.

?????????, ??? -c ?????????? ??????? ? **FTP** ? **HTTP**
????????????, ?????????? ?????????????????? ?????????????
"**Range**" (?.?. ?????????).

--progress=type

??? ?????????????? ??????????????. ?????????????? ?????????????: "**dot**" ? "**bar**".

"**bar**" ?????????????????? ?? ??????????????. ?????????? **--progress=bar**
???????????? ? ?????????????? ?????????????
????????????????? ?? ????????????? **ASCII**
(aka "????????????").1;5~ ????? ?????????????????? ?????? ?? TTY, ??????
????????????????? "т".Sp ?????????? **--progress=dot** ?????? ?????????????????? ?? ??? "**dot**
". ?????????????? ??????
????????????????? ?????????????????? ? ?????????? ?????? ??? ?????? ??????, ?????????? ?????????????
????
????????????????? ?????????????????? ?????????.

??? ?????????????????????? ??????? ??????, ?? ?????????? ?????????? ??? ?????? - **dot:style**

. ?????? ??????
"**default**
", ?????????? ?????????? ?????? ?????????????????? 1 K?, 10 ?????????? ? ?????????? ? 50 ? ???????
.
?????? "**binary**
" ?????? ?????? "?????????????????" ??? - 8 K? ??????????, 16 ?????????? ? ??????????
? 48 ?????????? ? ??????? (????????????????? ??????? ?? 384 K?). ?????? "**mega**
" ?????????????????? ???
????????????? ?????????? ??????? - ?????????? ?????????? ?????????????????? 64 K?, 8 ?????????? ? ??????
???? ?
48 ?????????? ? ??????? (????????????????? 3 M? ?? ?????????).

?? ?????????? ?????????????????? ??????? ?? ?????????????? ?????????????? ?????????? "**progress**" ? **.wgetrc**
. ??????
?? ?????????, ?????? ??? ?????????????? "**bar**
" ?????????????????????? ?????????? (? ?? ?????????? ??? ?????????? ?
stdout), ?????????? **--progress=bar:force**.

-N

--timestamping

????????????? ?????????????? ?? ?????.

-S

--server-response

????????????????? ??????????????, ?????????????? **HTTP** ?????????????? ? ??????????, ?????????????? **FTP**



?????????

--spider

????????? ?????????? Wget ??? ?????, ?. ?. Wget ?? ????? ?????????? ??????, ? ???
 ??? ??????
 ?????????? ?? ?????????. ??? ?????? ?????????? ??????????, ??????? ??????. ??????????:

wget --spider --force-html -i bookmarks.html

Wget ?? ?????????? ??? ?????????????? ????????????? WWW ???????.

-T seconds

--timeout=seconds

????? ?????????? ? ??????????. ?????? ?????????? ?? ?????????? 900 ?????? (15 ??????).
 ?????????? 0 ?????????? ?????????? ?????????? ??????????.

?????????????, ?? ?????????? ?????????? ?????????? ?????????? ?????? ?? ?? ???????, ??? ??????
 ??.

-w seconds

--wait=seconds

????? ? ?????????? ?????? ?????????????? ?????????????? (? ??.?. ?????????????). ??? ??????????
 ?????????????????? ??????????. ?????? ?????????? ?????????? ? ??????????, ?????????????? "m", ?
 ?????? - "h", ? ????? - "d"
 " ?????? ??????. ?????????? ?????????? ?????????? ?????? ??????????
 ?????????? ?????? ?????? ?????????????? (????????? ?????? dial up ??????).

--waitretry=seconds

????????????????? ?????? ??????? ?????? ?????????????? ??????????, ?????????? ??????????????. Wget
 ?????? ?????? 1 ??????????
 ?????? ?????????? ???????, 2 ?????????? ?????? ?????????? ?????????? ?????????? ?????? ?? ?????, ?
 ?. ?. ?? ?????????????,
 ?????????? ?????????????????? ? ??????????. ??????????, ??? ?????????????? ?????????????? 10, W
 get ?????? ?????? ? ??????
 (1 + 2 + ... + 10) = 55 ?????????? ??? ?????????? ??????.

??? ?????????????? ?????????????????? ?? ?????????????? ? ?????? wgetrc.

--random-wait

????????????? ??????????, ?????????? ?????????????????? ???-????????? ? ?????????? ?????????? ?????????,
 ?????? ??????????????????
 ?????????????????? ?????????? ????????? - ?????????????????? ??????????, ?????????, ??? Wget. ????? ??
 ?????????? ??????????????????????
 ?????? ?????? ?????????????? ?????????? ? ?????????? ?????????, ?????????????????????? ?? 0 ?? 2 * wa
 it (????????), ??? wait
 ?????????? ?????????????????? -w ??? ?????????????????? Wget.

????????? ?????????????? ??? ?????????????? ?????????????? Wget ? ?????? ??? ?????????????????? ?????????? ?????????
 ???.

-Y on/off

--proxy=on/off

????????????? ?????????-?????????. ?????????? ?? ?????????????? ?????? ?????????-????????? ??????????????.

-Q quota

--quota=quota

?????? ?? ??????? ?????????????????? ??????????. ?????????????????? ? ????????? (?? ??????????????), ?? (?
 ?????? k)



```

??? ? ?? (? ????? m).

??? ?????????? ?????, ??????? ????? ?????????????? ?? ?????, ??????? ??? ?? ?????????
??
????????? ?????? ??????. ?????????, ??? ? ? ?????? wget -Q10k
ftp://wuarhive.wustl.edu/ls-lR.gz [2],
??? ls-lR.gz ????? ?????????? ??????????. ????? ? ? ?????????? ? ?????????? ?????
?????
????????????? ??????????, ? ??????? ? ???????, ?????????? ? ?????, ??? ? ? ???????
??? ??????????.

????????? 0 ? ? inf ??????? ??????.

?????????? ?????????? ???????????

-nd
--no-directories
? ? ?????????? ?????????? ?????????? ? ? ?????????????? ??????????. ? ? ?????????? ?????
????, ? ? ?????
????? ?????????????? ? ? ? ? ??????. ? ? ? ? ? ?????? ?????? ? ? ??????????????, ? ?
????? ??????????
?? ? ?????? ??????????.n.

-x
--force-directories
????????????????????? ?????????? -nd
- ?????????? ?????????? ?????????? ??????? ? ??????? ?????????
?????????. ?????????? ?????????? ?????? wget -x http://fly.srk.fer.hr/robots.txt [3]
????????? ? ??????????
????????? fly.srk.fer.hr.

-nH
--no-host-directories
? ? ?????????? ?????????? c ?????? ?????? ? ??????????. ? ? ?????????? wget -r
http://fly.srk.fer.hr/ [4]
????????? ? ?????????? ?????????? ?????????? ?????????? ? fly.srk.fer.hr/
. ? ? ? ?????????? ??????????????
?????????? ?????????? ????????????.

--cut-dirs=number
????????????????? number
????????????? ??????? ??????. ?????????????????? ? ? ?????????? ?????????? ? ?????? ???????
?
?????????.

????????, ?????????, ??????? ftp://ftp.xemacs.org/pub/xemacs/
. ? ? ? ? ? ?????????? ? ? ? ? ?????????????? -r,
? ? ?????????????? ? ? ??????? ftp.xemacs.org/pub/xemacs/. ? ? ? ?????????? -nH
? ? ?????? ?????????????? ??????????
????? ?????????? ftp.xemacs.org/, ? ? ? ?????? ?????????????? pub/xemacs. ? ??????????
--cut-dirs ???????
number ??????????????. ?????????? ??????? ?????????????? --cut-dirs.

```

```

Без параметров -> ftp.xemacs.org/pub/xemacs/
-nH -> pub/xemacs/
-nH --cut-dirs=1 -> xemacs/
-nH --cut-dirs=2 -> .
--cut-dirs=1 -> ftp.xemacs.org/xemacs/

```



???? ?? ?????? ?????? ??????????? ?? ?????????? ??????????, ?? ?????? ?????????? ??
?? ?????????? ?? **-nd** ?
-P. ? ?????????? ?? **-nd**, **--cut-dirs** ?????????? ? ?????????????????? - ?????????? ???
-nH --cut-dirs=1, ?????????????
beta/ ?????????? ? xemacs/beta.

-P prefix

--directory-prefix=prefix

????????????? ?????????? ?????????, ? ????????? ?????? ????????????? ?????????? ?????????? ??
??? (??? ?????? ??????).
?? ????????????? ????? ?????????? ?????? . (????????? ??????????).

???????????? HTTP

-E

--html-extension

????? ??? ?????????????? ?????? **text/html** ? ??? ?????? ?? ?????????????????? ?? \
.[Hh][Tt][Mm][Ll]?, ??? ???????
???????????? ? ??? ?????? ?????? ????????????? **.html**
. ??? ?????? ????? ?????????? ??? ?????????????? ?????????? ??????????
.asp
, ?? ?? ?? ???????, ?????? ??? ?????????????? ? ??? ??????? Apache. ?????? ?????? ??
?????????? ??????
???????????? - ?????????? ?????????-????????? CGI ??????????. ?????????? ? URL ?????
http://site.com/article.cgi?25
?????? ?????????????? ??? *article.cgi?25.html*.

?????????????. ??? ?????????????? ??? ??????? ?????????????????? ?????????? ? ??????? ??????????????
, ?????????????? ??????
????????????????? ??????? ? ?????? ???????, ?. ?. Wget ?? ?????? ???????, ?????? ?? ???????
?? ?????????????? ?????
X.html ? ?????????????????? ? URL X. ?????? ?????????????? ?????????? ??????????????????, ??????????????????
?????? **-k** ? **-K**. ???
????? ?????????????????? ?????????? ?????????? ?????? ?????? ?????????????? ??? X.orig.

--http-user=user

--http-passwd=password

??? ?????????????????? user ? ??????? password ??? HTTP ??????????. ? ?????????????????? ?? ???
? ??????????, Wget ??????
????????????????? **"basic"** (?????????????????) ??? **"digest"** ??????????????????.

?????? ?????? ?????????????? ??? ?????????????????? ? ????????? ? ? ?????? URL.

-C on/off

--cache=on/off

?????????? ??? ?????????????? ?????????????????? ?? ?????????? ??????????. ??? ?????? Wget ??????????
????????????????????? ??????????
(**Pragma: no-cache**
). ?????? ?????????????????? ??? ?????????????? ?????????????? ?????????? ? ?????????-?????????.

?? ?????????????? ?????????????????? ??????????????.

--cookies=on/off

?????????? ??? ?????????????? ?????????????????? cookie. ??????? ??????????? ?????????? cookie ??
????????? ??????????????
"Set-Cookie
" ? ?????????? ?????????????? ??? ?? cookie. ?????????????? ??????? ?????????? ??????? ??????? ??????????????



??
?????????????. ?? ?????????? cookie ??????????????, ?? ?????? ?? ?? ???? ???????????.

--load-cookies file

?????????? cookie ?? file ?????? ?????? ??????????? HTTP. file ?????? ??????????? ???
???, ??? cookies.txt ?
Netscape.

???? ?????????? ??????????????? ??? ?????????????? ??????????. ??? ?????? Wget ?????????? ??
?? cookies, ?????????
????????? ??? ??????????????? ??? ?????????????? ? HTTP ??????????. ??? ?????????????? ??????
? ?????????????? - ???????
????????? Wget ????? ? cookies.txt. ?????? ?????????????????? ?????? cookie ? ?????? ??
?????????:

Netscape 4.x.
???? ??????????? ? ~/.netscape/cookies.txt.

Mozilla ? Netscape 6.x.
Mozilla ?????? cookies ? cookies.txt, ?????????????????? ???-?? ? ~/.mozilla
, ? ?????????? ???????
?????????. ?????? ???? ?????? ?????????????????? ???-?? ??????
~/.mozilla/default/some-weird-
string/cookies.txt.

Internet Explorer.
????? ?????????????????? cookie ??? Wget, ?????????? ???? | ?????? ? ?????????, ? ?????
??? ??????????
????????? ??????? cookie. ?????????????? ? Internet Explorer 5; ?????????? ?? ?????? ??
???????? ? ???????
?????????.

?????? ??????????????????
?????????? **--load-cookies**
?????? ?????????? ? cookie ? ?????????? Netscape, ?????????? ??????????????????? Wget.

???? ?? ?? ??????? ?????????????????? ?????????? --load-cookies, ????? ??????. ???? ??? ?
????????????????
?????????????????

wget --cookies=off --header "Cookie: I<name>=I<value>"

--save-cookies file

?????????? cookie ?? file ? ?????? ??????. ?????????????? cookie ?? ???????????????.

--ignore-length

?????????? HTTP ?????????? (???????? CGI ?????????) ?????????? ??????????? "Content-Length"
", ?????????? ?????????????
Wget, ??? ?????????????? ??? ?? ????. ? Wget ?????????????? ???? ?????????????? ??????????????
? ?????? ??????????????, Wget ?????? ?????????????????? ??????????? "Content-Length".

--header=additional-header

?????????????? **additional-header**
, ?????????????????? HTTP ??????????. ?? ??????? ?????????????? : ? ?????????? ?????? ?????.

?? ??????? ?????????????? ?????????????? ?????????????????????? ?????????????? ?????????????????????? --head
er ?????????????? ???.



```
wget --header="Accept-Charset: iso-8859-2" \  
--header="Accept-Language: hr" \  
http://fly.srk.fer.hr/ [4]
```

```
????????? ?????? ?????? ? ?????????? ?????????? ?????????? ??? ??????????????? ??????????  
????? ?? ??????  
????????????.
```

```
--proxy-user=user  
--proxy-passwd=password  
????????????? ??? ?????????????? user  
? ??????? ??? ?????????????? ?? ??????-?????????. ??????  
????????????????? ??? ?????????????? "basic".
```

```
--referer=url  
????????????? ??????????? `Referer: url  
' ? ?????? HTTP. ??????????????? ??? ?????????? ??????????, ??????????  
?????????????? ??????????? ?????????, ????? ????????? ?????????, ? ?????? ??????????? ?? ?????????.
```

```
-s  
--save-headers  
????????????? ?????????????, ????????????? HTTP ???????????.
```

```
-U agent-string  
--user-agent=agent-string  
????????????????????????????? ??? agent-string ??? ?????????? ?? HTTP ?????????.
```

```
?????????? HTTP ??????????? ?????????????? ????? ?????????????????????? ??????????? ??????????. Wget  
?? ??????????????  
????????????????????????? ??? Wget/version, ??? version - ??? ????????? Wget.
```

```
????????????? ?????????? ?????????? ?????????????? ?????????????? ?????????? ??? ?????????????????????, ??????????  
????????????????????? ???  
"Mozilla" ??? Microsoft "Internet Explorer  
". ?????? ?????????????? ?????????????? ??????????? ?????????? ???????????.
```

????????????? FTP

```
-nr  
--dont-remove-listing  
?? ?????????? ?????????????? ?????? .listing  
, ?????????????????? ??? ??????????? ?? FTP. ??? ?????? ?????????????? ??????????????  
? ?????????????? FTP ????????????. ?? ?????????????? ??????????? ??? ?????????? ?????????????????? ??????????????  
?????????????? ?????????? (?.  
?. ??????????????., ??? ?????? ??????????? ?????????????? ???????????).
```

```
?? ?????????? .listing  
, ?????????? ? ?????? ??????????????????. ??????????, ? ?????? ?????????? ?????? ??????????  
????????????????????? ?????????? ?? /etc/passwd ??? ???-?? ???.
```

```
-g on/off  
--glob=on/off  
????????????? ??? ?????????????? ?????????????? ?????????????????? ?????????????????? ?????????? (??????) ??  
FTP ??????????????. ???  
?????? ?????? *, ?, [ ? ]. ??????????:
```



wget ftp://gnjilux.srk.fer.hr/*.msg [5]

?? ?????????? ?????????????? ?????????????? ?????????? ??????????, ??? URL ?????????
????? ?????????.

?? ?????? ?????? ?????? URL ? ??????????. ??? ?????????? ??????? ?? ?????????? Unix FTP
(? ?????????????? ??????
Unix "ls").

--passive-ftp

????????? ?????????? ?????? FTP, ?????? ?????????????? ?????????????? ??????????. ??????????
????? ?? ?????????? fire?
wall.

--retr-symlinks

??? ?????????????? ?????????? ?????????? FTP, ??????, ?? ?????????? ?????????????? ??????????????
??? ??????, ??
?????????????. ??? ?????????? ?????????????? ???.

????????? **--retr-symlinks** ?????????? ?????????? ?????????? ??? ?????????, ?? ??? ?????????????.

????????, ??? ??? ?????????? ?? ?????????? ??? ?????????? ?????????????? ?????????.

????????? ?????????????? ??????????

-r

--recursive

????????? ?????????????? ??????????.

-l depth

--level=depth

????????????? ?????????? ?????????????? ?????????? **depth**. ?? ?????????? ?????????? 5.

--delete-after

????????? ?????????? ?????????? (?????????) ?????? ?? ??????????. ?????????????? ??? ??????????
??? ?????? ?????????? ??????
????????????????? ?????????? ?? ??????. ??????????:

wget -r -nd --delete-after <http://whatever.com/~popular/page/> [6]

????????? **-r** ?????????? ?????????? ?? ??????????, ?????????? **-nd**
????????????? ?????????? ??????????.

??? ?????????????? ?????????????? **--delete-after**, **--convert-links** ?????????????????.

-k

--convert-links

????? ?????????????? ?????????? ?????????????????????? ?????????? ? ?????????????? ??? ?????????????? ? ??
????????????? ??????????. ???
????????????? ?? ?????????? ?????????? ?????????? ?? ?????????? ??????????????, ? ?????????? ?? ??? ??????????
????????????? ?????????.

????????? ?????????? ?????????????? ?????????? ?? ?????? ?????????????:

- Ссылки на файлы, загруженные Wget изменяются на соответствующие относительные ссылки.
Например: если загруженный файл /foo/doc.html, ссылка на файл /bar/img.gif,



также загруженный, в нём будет выглядеть как ../bar/img.gif. Этот способ работает, если есть видимое соотношение между каталогами одного и другого файла.

- Ссылки на файлы, не загруженные Wget будут изменены на абсолютные адреса этих файлов на удалённом сервере.

Например: если загруженный файл /foo/doc.html содержит ссылку на /bar/img.gif (или на ../bar/img.gif), ссылка в файле doc.html изменится на <http://host/bar/img.gif> [7].

????????? ?????, ????????? ?????????? ?????????? ?????, ??????: ??? ???? , ?? ???
 ??? ???? ??????
 ?????????, ?????? ????? ?????????? ? ???? , ??? ? - ?????? ?????? ?????????? ?
 ??? ????? ? ??????????
 (???? ????? ?????????????). ??? ?????????????????? ??????, ?????????????? ??????????????
 ? ??????, ?????? ??
 ?????? ????????????? ?????????????? ???? ? ?????? ?????????, ?? ?????? ??? ???????????.

?????? ?????? ????????????? ?????????? Wget ?????, ?????? ?????? ??? ??????????. ??????
 ?????????, ??? ??????????
 -k ????????????? ????????????? ?????? ? ????????????? ??????????.

-K

--backup-converted

???????????????????? ?????? ????????? - ????????? ????????????? .orig. ????????? ?????????? ?
 ????? -N.

-m

--mirror

????????? ?????????? ??? ?????????????? ?????????? ??????. ??? ?????????? ?????? ??????
 ??? ??????????????: **-r -N**

-l inf -nr

. ??? ?????????????????? ?????????? ?????????????? ?????? ?????? ? ?????? ????????????????? ?
 ??????
 ??????????.

-p

--page-requisites

?????????? ??? ?????, ????????? ?????? ??? ?????????????? HTML ??????????. ?????????? ?????
 ???, ????, ?????????????
 ??????.

?? ????????????? ?????? ?????? ?? ??????????????. ?????????? **-r ? -l**
 , ????????????? ?????? ?????? ??????, ?? ?. ?.

Wget ?? ????????????? ?????????? ? ?????????????? ?????????????, ??? ??????????, ??? ??????????????
 ??? ?????????????.

?????????, 1.html ?????????? ??? "", ?? ????????? ?? 1.gif, ? ??? "<A>", ??????
 ?????? ?? ??????????

????????? 2.html. ?????????? 2.html ?????????????, ?? ?? ????????? - 2.gif ? ??????????
 ??? ?? 3.html.

??????, ??? ?????????????????? ?? ?????????????????? ??????.

????? ?????? ?????? ??????????:



wget -r -l 2 <http://> [8]<site>/1.html

?? 1.html, 1.gif, 2.html, 2.gif ? 3.html ??????????. ??? ?????, 3.html ??? 3.g
if, ?. ?. Wget ??????
??????? ?????? ?????????, ?? ????????? ?? ?????????, ????????? ?? 2 ? ??????????????????
? ??? ??????????????:

wget -r -l 2 -p <http://> [8]<site>/1.html

??? ?????? and ????????? 3.gif ?????????? 3.html ??????????????. ?????????????

wget -r -l 1 -p <http://> [8]<site>/1.html

?????????? ? ?????????? 1.html, 1.gif, 2.html ? 2.gif. ?????? ????????????? ????? ???????
??? HTML ?????????? ??
????? ?? ?????????????, ??????? ?? ?????????????? -r ? -l:

wget -p <http://> [8]<site>/1.html

??? ?????? Wget ?????? ?????? ?????? ??? ??? ????????????? -r, ?? ?????? ????????????? ???????
?? ? ?? ??????????????????????
??????. ?????? ?? ?????????, ?????? ?????????????????????? ?????? ?? ????????? ?????????? (??.? ?
??? ?????????????? ???????)
????? ?????????????, ??????????????:

wget -E -H -k -K -p <http://> [8]<site>/l<document>

? ? ?????????????, ?????? ?????????, ??? ??? Wget ?????????? ??????? - ??? URL, ??????????
? ? ?????? "<A>",
"<AREA>" ? "<LINK>", ?????? "<LINK REL="stylesheet">".

???????????? ?????????/????????????? ?????????????? ???????????

-A acclist --accept acclist

-R rejlist --reject rejlist

???????? ?????? ???????, ?????????????? ???????????, ?????????? ?????????? ??? ? ? ?????????? ??????
?????. ??????????????
????????? ?????? ??????? ?? ??????.

-D domain-list

--domains=domain-list

???????? ?????????? **domain-list**
, ? ?????????? ?????????????? ?????????????? ??????. ?????????????????? ??????????????. ?????? ??????????
?? ?????????? **-H**.

--exclude-domains domain-list

????????? ?????????, ? ?????????? ?? ?????????????? ?????????????? ??????

--follow-ftp

????????????? ?? ?????????? FTP ? HTML ??????????. ??????, ?????????? ?? ?????????? ?? ?????????????? F
TP ??????????????????

--follow-tags=list



Wget ?????? ???????????? ???????? HTML ??????, ? ???????? ????? ??????? ????
?. ?? ??????? ???????
????????????????? ????? ? ?????????????? ?????????? ??????? list ? ????? ???????????.

-G list

--ignore-tags=list

????????? **--follow-tags**

. ??? ?????????? HTML ?????? ??? ?????????????? ??????????, ????????? ? ? ??????????????
????????? ??????? list.

????????? ?????????? **-G**

??? ??????? ??? ?????????? ?????????? ????????? ? ? ? ?????????????????????? ??????????. ??
???????? ??????????????, ??? ??? ?????, ??????? ?????????

wget -Ga,area -H -k -K -r <http://> [8]<site>/l<document>

?? ???????, ??????? ?????????????? ??? ?????????? ?????? ?????????? ?????????? ??????????
--page-requisites.

-H

--span-hosts

?????????? ?????????? ?????? ?????????, ?? ????????? ?????? ???????.

-L

--relative

?????????? ??????? ? ? ?????????????????? ??????????. ??? ??? ? ?????????? ?????? ? ??????? ?
????????? ??????????????
?????? ? ? ??????.

-I list

--include-directories=list

??????, ?????????????? ??????????, ??????????, ?? ????????? ?????????? ?????????? ??????.
????????? ??????? list
?????? ?????????????? ?????????????????? ?????????.

-X list

--exclude-directories=list

??????, ?????????????? ??????????, ?????????????? ?????????? ??? ?????????? (?? . ??????????
??? ?? ??????????????)
?????????? ??????? list ?????? ?????????? ?????????????????? ?????????.

-np

--no-parent

?? ?????????????? ?????? ?????????????? ??????? ??? ?????????????? ??????????.

?????????

????????? ?????????????? ?? ??? ?????????????? ?????????? ?? ?????????????.

????????e ??????????????????

- Если вы вводите `wget http://физик.фед.ру/ [4]`
- Но что же будет, если соединение медленное, а файл длинный? Есть возможность обрыва связи перед завершением загрузки. В этом случае Wget будет продолжать попытки нового



соединения пока не кончится

число попыток --tries=45 <http://fly.srk.fer.hr/jpg/flyweb.jpg> [9]

- Теперь оставим Wget работать на фоне, а его сообщения писать в журнал log. Долго набирать --tries, так что wget --tries=45 -o log <http://fly.srk.fer.hr/jpg/flyweb.jpg> [9] & Символ амперсанда в конце указывает командному интерпретатору продолжать работу с пользователем, не дожидаясь завершения работы Wget. Чтобы программа делала повторы бесконечно используйте -t inf.

- Используйте FTP также просто wget берёт [10]

- Если вы укажете адрес каталога, Wget загрузит список содержимого этого каталога и конвертирует его в формат HTML <http://www.mit.edu/pub/gnu/> [11] **links index.html**

?????????? ?

- Если у Вас есть файл с URL, которые вы хотите загрузить, используйте параметр wget -i <file> Если вы укажете - вместо имени файла, URL будут читаться из стандартного ввода (stdin).

- Создайте копию с wget -r <http://www.gnu.org/> [12] -o gnu.log

- Как и выше, но конвертируя ссылки в HTML файлах в локальные для автономного просмотра wget -p -o convert-links -r <http://www.gnu.org/> [12] -o gnu.log

- Загрузить одну HTML страницу и все файлы, требуемые для отображения последней (напр. рисунки, файлы каскадных стилей). Так wget -p -o convert-links -r <http://www.server.com/dir/page.html> [13]

HTML страница будет сохранена в <http://www.server.com/dir/page.html> [13] и рисунки, каскадные стили и прочее будет сохранено в каталоге <http://www.server.com/> [14], кроме случая, когда файлы будут загружаться с других серверов.

- Как и выше, но без каталога <http://www.server.com/> [14]. Также все файлы будут сохранены в каталоге wget -p -o convert-links -r -nd -Pdownload \ <http://www.server.com/dir/page.html> [13]

- Загрузить <http://www.gnu.org/> [16]

- Сохранить <http://www.gnu.org/> [16]



more index.html

- Загрузить архивы в FTP-сервер <http://www.archive.wustl.edu/> [17].
- Загрузить из каталога GIF-файлы на HTTP сервере. Команда `wget http://www.server.com/dir/*.gif` [18] не будет работать, так как маскировочные символы не поддерживаются при загрузке по протоколу HTTP. Используйте `wget -r -l1 --no-parent -A.gif http://www.server.com/dir/` [19]. `-r -l1` включает рекурсивную загрузку с максимальной глубиной в 1. `--no-parent` выключает следование по ссылкам в родительский каталог, имеющую верхний уровень, `-A.gif` разрешает загружать только файлы с расширением .GIF. `-A "*.gif"` также будет работать.
- Предположим во время рекурсивной загрузки у вас нужно было срочно выключить/перезагрузить компьютер. Чтобы не потерять уже имеющиеся файлы, используйте: `wget -r -l1 http://www.gnu.org/` [12].
- Если вы хотите указать имя пользователя и пароль для HTTP или FTP сервера, используйте соответствующий синтаксис `wget -u[user] -P[password] URL[.dot] server[.dot] com/.emacs` `>ftp://hnksic:mypassword[at] unix[.dot] server[.dot] com/.emacs` [20].
- Вы хотите, чтобы загружаемые документы шли в стандартный вывод, а не в файлы? `wget -O - http://jagor.srce.hr/` [21] `http://www.srce.hr/` [22]. Если вы хотите устроить конвейер и загрузить все сайты, ссылки на которые указаны на одной странице: `wget -O - http://cool.list.com/` [23] | `wget --force-html -i -`

????????????????? ??????????????

- Для хранения зеркала страницы (или FTP каталога), используйте `--mirror (-m)`, что заменяет `-r -l inf -N`. Вы можете добавить Wget в `crontab` с запросом на проверку обновления каждое воскресенье:

```
0 0 * * 0 wget --mirror http://www.gnu.org/ [12] -o /home/me/weeklog
```

- ?? ????? ??????, ????? ?????????????????? ? ??????????. ?? ????? ?????????? ?????? ??????????????, ?? ???????, ??? ??? ??? ? ? ????? ?????????? ?????????? ?? ??????????. ??????? Wget ??? ???? ?????????? ?????? HTML ?????? ?????? ??????????????. ?????????:

```
wget --mirror --convert-links --backup-converted \
http://www.gnu.org/ [12] -o /home/me/weeklog
```

- ? ??? ? ? ?????????? ?????????? ?????????? HTML ?????? ? ??????????????, ?????????? ? . html, ??????????



index.cgi, ????? ????????? ????????? ? ? ????????????????????? ????? ?????? ?????? (content
-type = text/html) ?
???.html.

**wget --mirror --convert-links --backup-converted \
--html-extension -o /home/me/weeklog \
<http://www.gnu.org/> [12]**

? ?????????? ?????????? ?????????:

**wget -m -k -K -E <http://www.gnu.org/> [12] -o /home/me/weeklog
wget -m -k -K -E <http://www.gnu.org/> [12] -o /home/me/weeklog**

??????

/usr/local/etc/wgetrc
????????????????? ? ? ?????????? ?????????????? ?????? ??????????.

.wgetrc
????????????????????? ???? ??????????.

??????

? ? ?????? ?????????????? ?????? ? ? ?????? ? GNU Wget ? <"bug-wget(at)gnu.org"> (
? ??????????????).

????? ?????????????:

1. ??????????, ? ? ?????????? ?????????? ?????????????????? ??????. ???? Wget "?????????
?", ? ? ???? ???????.
???? ?????????? Wget ? ?????????????????? ?????????????????, ? ? ??????. ???? ? ? ??????
? ?????????, ? ? ? ?
????????, ? ? ? ? ?????? ?????????? ? ?????? ?????, ? ? ?????? ???? ??????.

2. ?????????????? ?????????? ?????? ? ?????????????? ?????????????? ??????????.

? ? ?????????? ?????????????? **.wgetrc**
, ?????????????? ?????????? ??????????, ?????????? ? ?????? ? ?????? (??? ??????
??? ????) ?????? ??????????.

3. ?????????? Wget ? ?????????????? **-d**
? ?????????? ?????? (?? ? ? ????). ?????????? ?????????? ?????? ??????????????
?????? ? ?????? ?????????????.

4. ???? Wget "?????????", ?????????????????? ?????????????? ? ? ?????????????, ?????????? "
gdb `which wget` core" ?
???????? "where" ? ? ?????????? ?????????? ?????????????????.

?. ??????

GNU Info ??? wget.

??????
Hrvoje Niksic <[hniksic\(at\)arsdigita.com](mailto:hniksic@arsdigita.com)>.



???????

? . ?????????? <admin(at)program.net.ua>

????????? ??????

Copyright (c) 1996, 1997, 1998, 2000, 2001 Free Software Foundation, Inc.

Источник (получено 2025-03-31 06:54):

<http://muff.kiev.ua/content/wget-rukovodstvo-gnu-wget>

Ссылки:

- [1] <ftp://sunsite.doc.ic.ac.uk/ls-IR.Z>
- [2] <ftp://wuarchive.wustl.edu/ls-IR.gz>
- [3] <http://sysadmins.el.kg/go.php?http://fly.srk.fer.hr/robots.txt>
- [4] <http://sysadmins.el.kg/go.php?http://fly.srk.fer.hr/>
- [5] ftp://gnjilux.srk.fer.hr/*.msg
- [6] <http://sysadmins.el.kg/go.php?http://whatever.com/%7Epopular/page/>
- [7] <http://sysadmins.el.kg/go.php?http://host/bar/img.gif>
- [8] http://sysadmins.el.kg/go.php?http://
- [9] <http://sysadmins.el.kg/go.php?http://fly.srk.fer.hr/jpg/flyweb.jpg>
- [10] <ftp://gnjilux.srk.fer.hr/welcome.msg>
- [11] <ftp://prep.ai.mit.edu/pub/gnu/>
- [12] <http://sysadmins.el.kg/go.php?http://www.gnu.org/>
- [13] <http://sysadmins.el.kg/go.php?http://www.server.com/dir/page.html>
- [14] <http://sysadmins.el.kg/go.php?http://www.server.com/>
- [15] <http://sysadmins.el.kg/go.php?http://www.lycos.com>
- [16] <http://sysadmins.el.kg/go.php?http://www.lycos.com/>
- [17] <ftp://wuarchive.wustl.edu/>
- [18] http://sysadmins.el.kg/go.php?http://www.server.com/dir/*.gif
- [19] <http://sysadmins.el.kg/go.php?http://www.server.com/dir/>
- [20] <ftp://hnksic:><span class=
- [21] <http://sysadmins.el.kg/go.php?http://jagor.srce.hr/>
- [22] <http://sysadmins.el.kg/go.php?http://www.srce.hr/>
- [23] <http://sysadmins.el.kg/go.php?http://cool.list.com/>