

WGET CHEAT SHEET

```
wget -e robots=off --random-wait -r -p -k -o crawl_log
```

-e robots=off	-E
-m	--html-extension (appends .html)
--mirror	
-r	-U agent-string
--recursive	--user-agent=agent-string
-p	-A acclist
--page-requisites	--accept acclist (comma-separated extensions)
-k	
--convert-links	-R rejlist
	--reject rejlist (comma-separated extensions)
-l depth	
--level=depth (5 is maximum)	-D domain-list
	--domains=domain-list (domains to follow)
-o logfile	
--output-file=logfile	--exclude-domains domain-list
-i file	
--input-file=file	-H
	--span-hosts
--random-wait	
	-L
	--relative (follow only relative links)
-nd	
--no-directories	
	-np
-nH	--no-parent
--no-host-directories	