

nixCraft

Linux Tips, Hacks, Tutorials, And Ideas In Blog

Convert HTML Page To a PDF Using Open Source Tool [Linux / OS X / Windows]

by VIVEK GITE

on OCTOBER 3, 2012 *last updated* OCTOBER 4, 2012

in [COMMAND LINE HACKS](#), [HOWTO](#), [OPEN SOURCE](#)

Do you need a simple open source cross-platform command line tool that converts web pages and HTML to a PDF file? Look no further, try wkhtmltopdf.

From the project home page:



Simple shell utility to convert html to pdf using the webkit rendering engine, and qt. Searching the web, I have found several command line tools that allow you to convert a HTML-document to a PDF-document, however they all seem to use their own, and rather incomplete rendering engine, resulting in poor quality. Recently QT 4.4 was released with a WebKit widget (WebKit is the engine of Apples Safari, which is a fork of the KDE KHTML), and making a good tool became very easy.

Software features

1. Cross platform.
2. Open source.
3. Convert any web pages into PDF documents using webkit.
4. You can add headers and footers.
5. TOC generation.
6. Batch mode conversions.
7. Can run on Linux server with an XServer (the X11 client libs must be installed).
8. Can be directly used by PHP or Python via bindings to libwkhtmltox.

A note about Debian / Ubuntu Linux user

You can install wkhtmltopdf using apt-get command:

```
$ sudo apt-get install wkhtmltopdf
$ sudo ln -s /usr/bin/wkhtmltopdf /usr/local/bin/html2pdf
```

Sample outputs:

```
[sudo] password for vivek:
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following NEW packages will be installed:
  wkhtmltopdf
0 upgraded, 1 newly installed, 0 to remove and 10 not upgraded.
Need to get 116 kB of archives.
After this operation, 303 kB of additional disk space will be used.
Get:1 http://debian.osuosl.org/debian/ squeeze/main wkhtmltopdf amd64 0.9.9-1 |
Fetched 116 kB in 2s (49.4 kB/s)
Selecting previously deselected package wkhtmltopdf.
(Reading database ... 274164 files and directories currently installed.)
Unpacking wkhtmltopdf (from .../wkhtmltopdf_0.9.9-1_amd64.deb) ...
Processing triggers for man-db ...
Setting up wkhtmltopdf (0.9.9-1) ...
```

Download wkhtmltopdf

Visit this [page](#) to grab wkhtmltopdf for Linux / MS-Windows / Apple Mac OS X. You can also use the [wget command](#) as follows:

```
$ wget http://wkhtmltopdf.googlecode.com/files/wkhtmltopdf-0.11.0_rc1-static-amd64.tar.bz2
```

Sample outputs:

```
Resolving wkhtmltopdf.googlecode.com... 74.125.135.82, 2404:6800:4001:c01::52
Connecting to wkhtmltopdf.googlecode.com|74.125.135.82|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 11175276 (11M) [application/octet-stream]
Saving to: `wkhtmltopdf-0.11.0_rc1-static-amd64.tar.bz2'

100%[=====>] 1,11,75,276 480K/s in 23s

2012-10-04 01:21:43 (477 KB/s) - `wkhtmltopdf-0.11.0_rc1-static-amd64.tar.bz2' saved [11175276]
```

Install wkhtmltopdf under Linux

Type the following [tar command](#) to extract files:

```
$ tar xvf wkhtmltopdf-0.11.0_rc1-static-amd64.tar.bz2
```

Sample outputs:

```
wkhtmltopdf-amd64
```

Install the same in your private ~/bin/ directory or in /usr/local/bin directory:

```
$ mv wkhtmltopdf-amd64 ~/bin/  
ln -s ~/bin/wkhtmltopdf-amd64 ~/bin/html2pdf
```

OR

```
$ sudo mv wkhtmltopdf-amd64 /usr/local/bin/  
ln -s /usr/local/bin/wkhtmltopdf-amd64 /usr/local/bin/html2pdf
```

How do I use wkhtmltopdf?

The syntax is as follows:

```
html2pdf http://www.cyberciti.biz/path/to/url.html output.pdf  
html2pdf http://www.cyberciti.biz/blog/print/url-slug.html output.pdf  
html2pdf -option1 -option2 http://www.cyberciti.biz/blog/print/url-slug.html output.pdf
```

Example: Simple html to pdf file

In this example, convert out [bash for loop page](#) to a pdf file:

```
$ html2pdf http://www.cyberciti.biz/faq/bash-for-loop/print/  
/tmp/bash.for.loop.pdf
```

Sample outputs:

```
Loading pages (1/6)  
Counting pages (2/6)  
Resolving links (4/6)  
Loading headers and footers (5/6)  
Printing pages (6/6)  
Done
```

To view generated pdf file [click here](#). Verify pdf file, enter:

```
$ file /tmp/bash.for.loop.pdf
```

Sample outputs:

```
/tmp/bash.for.loop.pdf: PDF document, version 1.4
```

Use the [pdftinfo command](#) to prints the contents of the ‘Info’ dictionary (plus some other useful information) from a Portable Document Format (PDF) file:

```
$ pdftinfo /tmp/bash.for.loop.pdf
```

Sample outputs:

Title: Frequently Asked Questions About Linux / UNIX » Bash For Loop Example

Creator: wkhtmltopdf

Producer: wkhtmltopdf

CreationDate: Thu Oct 4 01:29:33 2012

Tagged: no

Pages: 4

Encrypted: no

Page size: 595 x 842 pts (A4)

File size: 98792 bytes

Optimized: no

PDF version: 1.4

Grayscale pdf

The following PDF will be generated in grayscale:

```
$ htmI2pdf -g http://www.cyberciti.biz/faq/bash-for-loop/print/bash.for.loop.pdf
```

Set orientation to Landscape or Portrait

Use the following syntax:

```
$ htmI2pdf -O Landscape http://www.cyberciti.biz/faq/bash-for-loop/print/bash.for.loop.pdf
```

Where,

- -O Landscape|Portrait. The default is Portrait.

How do I set page size?

Use the following syntax:

```
$ htmI2pdf -S SIZE http://www.cyberciti.biz/faq/bash-for-loop/print/bash.for.loop.pdf
```

Where,

- -s Size : Set paper size to: A4, Letter, etc. (default A4)

How do I generate table of content?

A table of content can be added to the document by adding a toc the command line option. For example:

```
$ htmI2pdf toc http://www.cyberciti.biz/faq/bash-for-loop/print/bash.for.loop.pdf
```

Sample outputs:

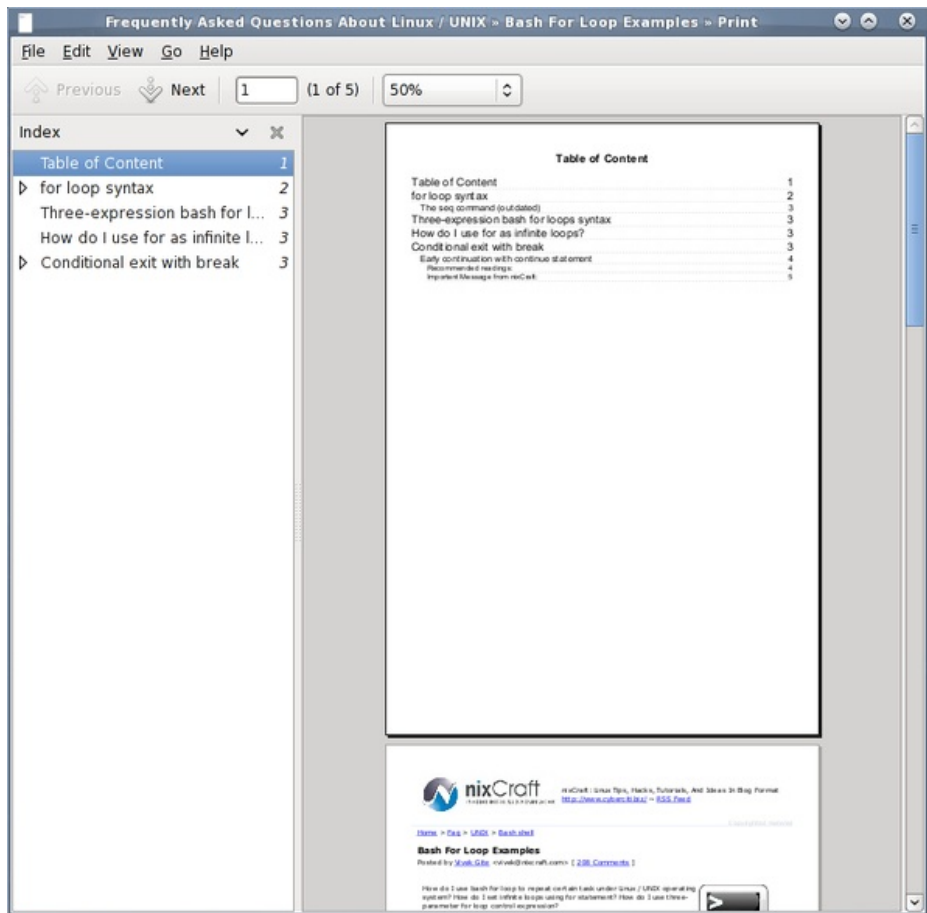


Fig.01: wkhtmltopdf in action

Please note that the table of content is generated based on the H tags in the input documents.

How do I see all available options?

You can see a list of commonly used options, enter:

```
$ wkhtmltopdf --help
```

OR see all available options i.e. display more extensive help, detailing less common command switches, run:

```
$ wkhtmltopdf -H | less
```

References:

- [wkhtmltopdf](http://www.nixcraft.com/bash-for-loop-examples.html) project home page.



- Twitter
- Facebook
- Google+
- Download PDF version
- Found an error/typo on this page?

More like this:

- [pinfo: Read Linux Info Documentation in Colors](#)
- [diff Command: Colorize Output On The Unix / Linux Command Line](#)
- [Linux / Unix pv Command: Monitor Progress of Data Sent Via a Pipe](#)
- [Apache / Nginx: Visualize Web Server Access Log In Real Time](#)
- [Goldendict: A Feature-rich Dictionary Lookup Linux Program](#)
- [Linux ifdata Command: See Network Interface Info Without Parsing ifconfig...](#)
- [Display Awesome Linux Logo With Basic Hardware Info Using screenfetch and...](#)
- [agedu: Unix / Linux Command For Tracking Down Wasted Disk Space](#)
- [Testing HTTP Status: 206 Partial Content and Range Requests](#)
- [Linux / Unix Desktop Fun: Cat And Mouse Chase All Over Your Screen](#)



Tagged as: [Apple](#), [Linux](#), [MS-Windows](#), [Unix](#), [wkhtmltopdf](#)

Comments on this entry are closed.

Next post: [HowTo: Linux Hard Disk Encryption With LUKS \[cryptsetup Command \]](#)

Previous post: [Linux / Unix Desktop Fun: Cat And Mouse Chase All Over Your Screen](#)

Samsung Galaxy J5

Mobil och abonnemang fr.
165 kr/mån Fria SMS,
MMS, samtal. Surf ingår.

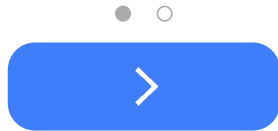


Advertise Here

To search, type and hit enter

Börja växa med Google

Utnyttja din kupong
värd 750 kronor Se
till att du syns med
AdWords



Advertise Here

FEATURED ARTICLES:

30 Cool Open Source Software I Discovered in 2013

30 Handy Bash Shell Aliases For Linux / Unix / Mac OS X

Top 30 Nmap Command Examples For Sys/Network Admins

25 PHP Security Best Practices For Sys Admins

20 Linux System Monitoring Tools Every SysAdmin Should Know

20 Linux Server Hardening Security Tips

Linux: 20 Iptables Examples For New SysAdmins

Top 20 OpenSSH Server Best Security Practices

Top 20 Nginx WebServer Best Security Practices

20 Examples: Make Sure Unix / Linux Configuration Files Are Free From Syntax Errors

15 Greatest Open Source Terminal Applications Of 2012

My 10 UNIX Command Line Mistakes

Top 10 Open Source Web-Based Project Management Software

Top 5 Email Client For Linux, Mac OS X, and Windows Users

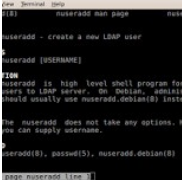
The Novice Guide To Buying A Linux Laptop



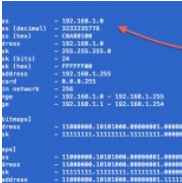
Open Source Photoshop Software



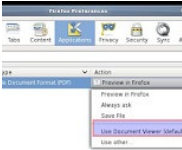
Howto Open a PDF file under Linux / FreeBSD



pdftotext: Linux / UNIX Convert a PDF File To T...



Print a Web Page From a Shell Prompt



pdftimages: Extract and Save Images From A Porta...



HowTo: Find Out My Linux Distribution Name and ...



Bash: Display Web Page Content In Terminal

HowTo: Linux / Unix


Don't Miss Any Linux Tips

Get nixCraft in your inbox. It's free:


Your email address

Subscribe

Follow Us



nixCraft

 Följ

+ 246 254

 Follow @nixcraft

54.8K followers