

No doubt you have been through this scenario time and time again. A friend or someone from the office sends you a CD list, club statistics or even a recipe with the message: "You probably use Excel, so I'm attaching the file in *.xls* format." Now, as a Linux user you might not use Excel at all (or not feel like rebooting), and you may not want to launch a processor beast like *StarOffice* – so you have a problem. Other aggravating situations may occur when you attempt to access statistics placed on the Web by local authorities or government, as they too will often rely on "standards" à la Redmond.

Sooner or later, when the pressure starts to build, you will probably start looking for a suitable tool – you could try **Freshmeat** for example, my personal favorite site for Open Source software. And that is where you will find *xlhtml*, a tool written by Charles Wyble that looks just the job for Excel plagued Linux users. The program converts Excel spreadsheets to **HTML**, allowing you to view them in a standard web browser.

## A Trip to Chicago

*xlhtml* maintains a homepage at <http://chicago.sourceforge.net/xlhtml/>. And that is the place to download the source archive that you will need to install the software. Other pre-requisites are of course the GNU C Compiler and the usual suspects such as **make** and the *glibc-dev* package.

The installation is a matter of a few simple steps:

```
tar xzf xlhtml-0.5.tgz
cd xlhtml-0.5
./configure
make
su (Input the root password)
make install ; exit
```

Instead of the last line you can also use the tool discussed in a recent "out of the box", *checkinstall* [1]:

## OUT OF THE BOX

There are thousands of tools and utilities for Linux. "Out of the box" takes a pick of the bunch and each month suggests a little program, which we feel is either absolutely indispensable or unduly ignored.

## Out of the box

# eXcellent

It has been quite a while since we introduced antiword – a filter for Word documents, in this column. A similarly useful piece of software for Excel spreadsheets has been sorely missed so far. Enter *xlhtml* to close that gap.

BY CHRISTIAN PERLE



Mail Veith, visipix.com

```
checkinstall ; exit
```

You can then use your distribution's package manager to remove the program, if required.

After completing these steps, you will discover that *xlhtml* has been placed in */usr/local/bin*, a directory that should be mapped in your **PATH** variable.

## Trial Run

You will need an Excel document to test *xlhtml*. If you happen not to have one available on your hard disk, you can use the *Test.xls* file from the *xlhtml* archive.

This spreadsheet uses all the special effects currently supported, for example colors, fonts and various attributes. Type the following to start your first test

```
xlhtml Test.xls > Test.html
```

and view the HTML file *Test.html* created by this command in a browser of your choice. Figures 1 (Excel for Windows) and 2 (the *xlhtml* view in Netscape) show a test performed with a simple Excel file which includes some cell errors.

But not only GUI web browsers can be used to view the spreadsheet created by

Month	Consumption (kWh)	Consumption (kWh)	%	Changes	%	Power(kWh)
January	0	**	**	0.00	**	**
February	0	**	**	0.00	**	**
March	0	**	**	0.00	**	**
April	0	**	**	0.00	**	**
May	0	**	**	0.00	**	**
June	0	**	**	0.00	**	**
July	0	**	**	0.00	**	**
August	0	**	**	0.00	**	**
September	0	**	**	0.00	**	**
October	0	**	**	0.00	**	**
November	0	**	**	0.00	**	**
December	0	**	**	0.00	**	**
Summary	**	**	**	**	**	**

Figure 1: The spreadsheet viewed on a hostile operating system

Month	Consumption (kWh)	Consumption (kWh)	%	Changes	%	Power(kWh)
January	0	**	**	0.00	**	**
February	0	**	**	0.00	**	**
March	0	**	**	0.00	**	**
April	0	**	**	0.00	**	**
May	0	**	**	0.00	**	**
June	0	**	**	0.00	**	**
July	0	**	**	0.00	**	**
August	0	**	**	0.00	**	**
September	0	**	**	0.00	**	**
October	0	**	**	0.00	**	**
November	0	**	**	0.00	**	**
December	0	**	**	0.00	**	**
Summary	**	**	**	**	**	**

Figure 2: The converted spreadsheet in Netscape

*xlhtml*. Even console based programs of the same ilk, such as *w3m* or *links* can show their prowess, when confronted with documents of this type (Figure 3).

## The cutting table and third-party formats

If you are only interested in a specific part of the spreadsheet and you know the row and column references, you can use the *-xr* (“extract row”) and the *-xc* (“extract column”) references to pass instructions on to the program. There is also an *-xp* (“extract page”) option that allows you to select specific pages in the document. To extract rows 2 through 5 and columns 0 through 2 from the *Book1.xls* spreadsheet

and view the results directly on the console, you can type the following (if you have installed *w3m*):

```
xlhtml -xr:2-5 -xc:0-2
2 Book1.xls | w3m -T
text/html
```

This command assumes that *xlhtml* will output the HTML page to **standard output** and the *w3m* browser will read directly from **standard input**. The pipe character (“|”) is used to redirect the output to the following command. You use the *-T* option to give *w3m* details on the data stream format – *text/html* refers to HTML in this case.

*Xlhtml* offers additional output format options. The *-xml* option converts Excel to XML (“Extensible Markup Language”), *-csv* creates “Comma Separated Values”, and the *-asc* option, creates pure ASCII text. The last two formats are only available with the *-x* options. If you want the content of the spreadsheet called *cdlist.xls* in a text-only format, you type:

```
xlhtml -asc -xp:99 cdlist.xls
```

Since we want to read all the document, the value for the *-xp* argument must be larger than the actual number of pages.

## Midnight Commander can do

Just like the Word filter *antiword* [2], we can use *xlhtml* as a filter for the built in file manager, *Midnight Commander (mc)*. To do so, you simply add the following two lines to the *~/.mc/bindings* file:

```
shell/.xls
View=%view{ascii} xlhtml %f
| w3m -T text/html -dump
```

To view Excel spreadsheets in *mc*, you simply select the file and press the [F3] key.

## Powerpoint

*xlhtml*'s author is a very hard-working guy: The source archive for *xlhtml* not only contains the Excel converter, but also a program called *phtml*. This can convert Powerpoint files to HTML. To call the program, you can use the following syntax:

```
phtml powerpoint_file.ppt >
html_file.html
```

But don't expect too much – the current version merely extracts the text.

Month	Consumption (kWh)	Consumption (kWh)	%	Changes	%	Power(kWh)
January	0	**	**	0.00	**	**
February	0	**	**	0.00	**	**
March	0	**	**	0.00	**	**
April	0	**	**	0.00	**	**
May	0	**	**	0.00	**	**
June	0	**	**	0.00	**	**
July	0	**	**	0.00	**	**
August	0	**	**	0.00	**	**
September	0	**	**	0.00	**	**
October	0	**	**	0.00	**	**
November	0	**	**	0.00	**	**
December	0	**	**	0.00	**	**
Summary	**	**	**	**	**	**

Figure 3: Spreadsheet view in *w3m* on a character based console

## GLOSSARY

**Freshmeat:** A major resource for current projects in the Open Source area, is at <http://freshmeat.net/> on the Web.

**HTML:** “HyperText Markup Language”, the markup language originally developed by CERN for World Wide Web pages. So-called Tagmark specific text passages as titles, lists, tables and so on.

**make:** A program for organizing source code compilation. The *makeconfiguration* file (*Makefile*) can contain information on dependencies for the individual program modules.

**PATH:** This variable consists of a list of directories separated by colons. The shell will search these directories for any commands typed without a path. Thus, *top* will be located in */usr/bin/top*.

**Standard input, standard output:** A large number of command line programs allow you to leave out the name of the input file. In this case the program reads from standard input, which is normally the keyboard. If you leave out the name of the standard output file, many programs will default to standard output, that is, display the output on your terminal. Output can be redirected into a file using the *>* character or to another command using the *|* character.

## INFO

[1] CheckInstall: “Say Hello Wave Goodbye”, Linux Magazine Issue 22, p78

[2] Antiword: “Against It!”, Linux Magazine Issue 15, p82

