Fun With The ‘AWK’ Command

Today, I’ll show you a real-world use of the venerable AWK command. I use the AWK command fairly regularly.

First, a little about AWK. The AWK command has been around since 1977 and is named after the three people who wrote the original version: Alfred **A**ho, Peter **W**einberger, and Brian **K**ernighan. It was first officially released in Version 7 Unix (V7) and has been a valuable text processing program ever since.

This is just one instance of how I used it to process some text.

I was going through the statistics at the Music For Us site. Among the statistics is a list of countries from which MFU has had visitors. Given the age of the site, it’s a fairly long list.

The format of the text was as follows:

[code]1 United States United States 8,645
2 Germany Germany 1,044
3 China China 846
4 France France 672
5 Canada Canada 538[/code]

I only cared about the name of the countries, but it was a very long list (more than 100 entries) and I didn’t want to edit them out manually. So, it was a pretty easy decision to use AWK to process it.

First, I copied the text into a document called ‘countries.txt’. I then opened my terminal and headed to that directory. This was the command I used:

[code]awk ‘{ print $2 }’ countries.txt > finished.txt[/code]

So, I basically told it to use AWK to process countries.txt, to extract the second column, and to write that data to a file called finished.txt. In this case, I could have easily extracted any of the columns without much difficulty. You simply need to change the $2 to whichever column you want.

If you'd like a full explanation of the command, you can actually check it out on one of my favorite sites, explainshell.com. (That link will take you directly to the explanation of this specific command.)

Anyhow, the output was as expected. It looks like:

[code]United
Germany
China
France
Canada[/code]

(If you want to see the full list, check out the paste here. By the way, you can feel free to use that site for your own paste needs. We have our own free pastebin service.)

I guess the final verdict is that Linux really lets me be lazy! Instead of processing that text manually, I used the AWK command and processed it all in less than a second. I didn't have to sit there and copy/paste. I didn't have to spend the time verifying it. I just had to run one simple command and I had the data I needed, nice and neat.

The thing is, I really only need the basics so I have giant holes in my knowledge – and yet, I know enough to do things like that, making my life easier. There's always so much more to learn, and that's awesome.