



UNIVERSIDAD CATOLICA
DE TEMUCO

XML con Perl yPython

Profesor: Alejandro Mellado Gatica
Carrera: Ingeniería Civil Informática
Alumno: Cristian Beltrán Concha

Introducción

Desde un archivo ODS que luego se transformo a CSV, se trabajo con el lenguaje Perl, leyendo y escribiendo un archivo XML con sus respectivas etiquetas. Luego con un script en Python leer el archivo XML, pero trabajandolo con una libreria para dicho archivo se espera obtener los datos de todos los registros que se insertaron desde el CSV con Perl.

Desarrollo

PERL

Con el archivo CSV se extraen los datos, primeramente las cabeceras para tener así las etiquetas del XML.

```
9  open FILE, "lista_de_usuarios.csv" or die $!;
10 open my $fh, '>', 'archivoXML.xml' or die $!;    # archivo final
11
12
13 my @cabeceras = split(',', <FILE>);    # obtiene las etiquetas
14
15 foreach my $v (@cabeceras)
16 {
17     $v =~ s/s/_/;    # reemplaza espacios, para evitar errores
18 }
19
```

Con estas líneas se agregan las cabeceras a un Array y luego se reemplazan los espacios en blanco que tenga cada etiqueta por un '_' para evitar errores en la estructura del XML.

Las siguientes líneas imprimen en el archivo final XML. Primeramente la cabecera y luego comienza con la inserción de los registros.

Trozea la línea del archivo CSV seccionándolo con el carácter ',' para cada dato.

Luego recorre cada dato del arreglo cabecera para las etiquetas de los datos del registro y lo agrega al archivo.

```
20 print $fh "<?xml version='1.0' encoding='UTF-8'?>\n";
21 print $fh "<Personas>\n";
22 my $j=1;
23 while(my $linea = <FILE>)
24 {
25     my @info = split(',', $linea);
26     print $fh "<Persona$j>\n";
27     my $i=0;
28     foreach my $dato (@info)
29     {
30         my $c = @cabeceras[$i];
31         print $fh "<$c>$dato</$c>\n";
32         $i++;
33     }
34     print $fh "</Persona$j>\n";
35     $j++;
36 }
37
38
```

Finaliza por agregar la ultima etiqueta y cierra los archivos e imprime que todo se realizo de manera correcta.

```
39 print $fh "</Personas>\n";
40
41 close($fh);
42 close(FILE);
43 print "Archivo Creado Exitosamente!!\n";
44
```

XML

El primer registro del XML queda de la siguiente manera, luego de ejecutar el script Perl.

```
1 <?xml version='1.0' encoding='UTF-8'?>
2 <Personas>
3 <Personal>
4 <RUT>19530300</RUT>
5 <DV>K</DV>
6 <REGISTRO>2015013473</REGISTRO>
7 <APELL.PATERNO>ACUÑA</APELL.PATERNO>
8 <APELL.MATERNO>VIVEROS</APELL.MATERNO>
9 <NOMBRES>JAIME ANDRES</NOMBRES>
10 <SEXO>MASCULINO</SEXO>
11 <FECHA_NACIMIENTO>4/30/1997</FECHA_NACIMIENTO>
12 <ESTADO>REGULAR</ESTADO>
13 <AÑO_INGRESO>2015</AÑO_INGRESO>
14 <COD.CARRERA>154</COD.CARRERA>
15 <NOMBRE_CARRERA>INGENIERÍA CIVIL EN INFORMÁTICA</NOMBRE_CARRERA>
16 <PLAN>2</PLAN>
17 <COD.MENCION_>0
18 </COD.MENCION_>
19 </Personal>
```

PYTHON

Desde python con su libreria para XML se trabaja con el archivo creado con Perl, imprimiendo los datos por pantalla.

```
9 import xml.etree.ElementTree as XML
10
11 root = XML.parse('archivoXML.xml').getroot()
12
13 for p in root:
14     for v in p:
15         print v.tag + ': ' + v.text
16     print '\n'
17
```

Las primeras dos lineas importamos la libreria y abrimos el archivo XML. Luego en el primer ciclo recorremos los registros y el segundo los datos de cada registro imprimiendolos en pantalla.

```
RUT: 19530300
DV: K
REGISTRO: 2015013473
APELL.PATERNO: ACUÑA
APELL.MATERNO: VIVEROS
NOMBRES: JAIME ANDRES
SEXO: MASCULINO
FECHA_NACIMIENTO: 4/30/1997
ESTADÔ: REGULAR
AÑO_INGRESO: 2015
COD.CARRERA: 154
NOMBRE_CARRERA: INGENIERÍA CIVIL EN INFORMÁTICA
PLAN: 2
COD.MENCION_: 0

RUT: 19530301
DV: 8
REGISTRO: 2015014029
APELL.PATERNO: ACUÑA
APELL.MATERNO: LARA
NOMBRES: RICHARD ALEX
SEXO: MASCULINO
FECHA_NACIMIENTO: 2/28/1997
ESTADÔ: REGULAR
AÑO_INGRESO: 2015
COD.CARRERA: 154
NOMBRE_CARRERA: INGENIERÍA CIVIL EN INFORMÁTICA
PLAN: 2
COD.MENCION_: 0
```

Ejecutando el script de Python vemos los registros con cada uno de sus datos.

Conclusión

El facil manejo de XML con python hace muy rapido el obtener los datos desde el archivo.

El escribir el archivo sin librerias desde Perl para luego leerlo con Python con su libreria nos muestra la ayuda que nos presenta un archivo con etiquetas como es XML para el manejo de los datos con distintas plataformas y lenguajes.

Referencias

<https://docs.python.org/3/library/xml.etree.elementtree.html>

<http://www.perlfect.com/articles/perlfile.shtml>