

How to convert XML file to Mathematica Grid and to Latex table

Nasser M. Abbasi

July 1, 2016 compiled on — Friday July 01, 2016 at 02:24 AM

This is an example how to read a plain text XML file in Mathematica, and display its content in a Mathematica Grid (table like).

I used an example XML I found on the net (ref [2]), which is a plain text file that contains some made up student information.

```
<?xml version="1.0" encoding="utf-8"?>
<students>

  <student>
    <studentnumber>284008</studentnumber>
    <firstname>Benjamin</firstname>
    <lastname>Carson</lastname>
    <dateofbirth>04/10/1995</dateofbirth>
    <Gender>2</Gender>
  </student>

  <student>
    <studentnumber>826084</studentnumber>
    <firstname>Gertrude</firstname>
    <lastname>Simms</lastname>
    <dateofbirth>8/22/1993</dateofbirth>
    <Gender>1</Gender>
  </student>

  <student>
    <studentnumber>628460</studentnumber>
    <firstname>Paul</firstname>
    <lastname>Sandt</lastname>
    <dateofbirth>12/24/1997</dateofbirth>
    <Gender>3</Gender>
  </student>

  <student>
    <studentnumber>792714</studentnumber>
    <firstname>Chrissie</firstname>
    <lastname>Burchs</lastname>
    <dateofbirth>02/06/1993</dateofbirth>
    <Gender>1</Gender>
  </student>

</students>
```

The first step is to download the above file to some folder, then create the following Mathematica code

```
SetDirectory[NotebookDirectory[]];
r = Import["data.xml", "XML"]
```

The above now displays the symbolic XML

```

XMLObject [Document] [{XMLObject [Declaration] [Version->1.0,
Encoding->utf-8]},
XMLElement [students, {},
{
XMLElement [student, {},
{XMLElement [studentnumber, {}, {284008}],
XMLElement [firstname, {}, {Benjamin}],
XMLElement [lastname, {}, {Carson}],
XMLElement [dateofbirth, {}, {04/10/1995}],
XMLElement [Gender, {}, {2}]}],
XMLElement [student, {},
{XMLElement [studentnumber, {}, {826084}],
XMLElement [firstname, {}, {Gertrude}],
XMLElement [lastname, {}, {Simms}],
XMLElement [dateofbirth, {}, {8/22/1993}],
XMLElement [Gender, {}, {1}]}],
XMLElement [student, {},
{XMLElement [studentnumber, {}, {628460}],
XMLElement [firstname, {}, {Paul}],
XMLElement [lastname, {}, {Sandt}],
XMLElement [dateofbirth, {}, {12/24/1997}],
XMLElement [Gender, {}, {3}]}],
XMLElement [student, {},
{XMLElement [studentnumber, {}, {792714}],
XMLElement [firstname, {}, {Chrissie}],
XMLElement [lastname, {}, {Burchs}],
XMLElement [dateofbirth, {}, {02/06/1993}],
XMLElement [Gender, {}, {1}]}]
}],
{}]

```

Now we read all the fields, using Cases

```

students=Cases[r, XMLElement["student", _, _], Infinity];
numbers=Flatten[Cases[students, XMLElement["studentnumber", _, x_]->x, Infinity], 1];
firstName=Flatten[Cases[students, XMLElement["firstname", _, x_]->x, Infinity], 1];
lastName=Flatten[Cases[students, XMLElement["lastname", _, x_]->x, Infinity], 1];
gender=Flatten[Cases[students, XMLElement["Gender", _, x_]->x, Infinity], 1];
dateofbirth=Flatten[Cases[students, XMLElement["dateofbirth", _, x_]->x, Infinity], 1];

```

Now put them in a grid

```

Grid[Join[{"first_name", "last_name", "gender", "DOB"},
Transpose[{firstName, lastName, gender, dateofbirth}]],
Frame -> All]

```

This is the result

first name	last name	gender	DOB
Benjamin	Carson	2	04/10/1995
Gertrude	Simms	1	8/22/1993
Paul	Sandt	3	12/24/1997
Chrissie	Burchs	1	02/06/1993

Finally, the output is converted to Latex in the fly, and compiled to HTML and pdf using tex4ht and pdfplatex. Here is the result

first name	last name	gender	DOB
Benjamin	Carson	2	04/10/1995
Gertrude	Simms	1	8/22/1993
Paul	Sandt	3	12/24/1997
Chrissie	Burchs	1	02/06/1993

This is the Mathematica code used to convert the table to Latex

```
str="\documentclass[11pt]{book}
\\begin{document}
\\begin{tabular}{";
Do[str=str<>"|1",{n,1,Length[students]}}];
str=str<>"|\n";
str=str<>"first_name&last_name&gender&DOB\\\\\n\\hline\n";
Do[
str=str<>firstName[[n]]<>"<"<>lastName[[n]]<>"<"<>gender[[n]]
<>"<"<>dateofbirth[[n]]<>"\\\\\n\\hline\n",
{n,1,Length[students]}}];
str=str<>"\\end{tabular}
\\end{document}
";

fileName = "output.tex";
If[FileExistsQ[fileName], DeleteFile[fileName]];
file = OpenWrite[fileName, PageWidth -> Infinity];
WriteString[file, str];
Close[file];
```

That is all.

References

1. <https://reference.wolfram.com/language/XML/tutorial/TransformingXML.html>
2. <http://www.functionx.com/xml/Lesson01.htm>