
gedcom55XMLtoXSL-FO.xsl

Chad Albers

August 8, 2008

1. Overview

gedcom55XMLtoXSL-FO.xsl is an Extensible Stylesheet Language¹ (XSL) stylesheet that uses XSLT² to transform a GEDCOM 5.5 XML³ document to a XSL-FO⁴ stylesheet. This stylesheet can be converted to a Portable Document Format⁵ (PDF) file using Apache's Formatting Objects Processor⁶ Java application, fop. The pdf document resembles genealogy record keeping forms produced by ProGenealogists⁷. They specifically resemble the forms called "Family Group Sheet Page 1"⁸ and "Family Group Sheet Page 2"⁹.

1.1. Stylesheet Parameters

Several parameters can be supplied to XSLT processors which affect the data included in the document and the document's "look and feel."

1.1.1. Output Data Parameters

- IncludeIDs - accepts either 'true' or 'false' and acts as a flag to include the XREF ID numbers of family record elements (FAM) and the individual family member elements (INDI), including the spouses of children
- IncludeDateGenerated - accepts either 'true' or 'false'. If 'true', the date that the XSLT processor applied this stylesheet to a GEDCOM 5.5 XML document is included in the footer of the pdf document. This parameter relies on the <http://exslt.org/dates-and-times> extension. If the XSLT processor does not support this extension, the stylesheet may fail to be applied to the XML document. To be on the safe side, this parameter defaults to 'false.'
- FamID - accepts the XREF ID of one FAM element. If this value is supplied, the stylesheet will produce a document containing only the family record whose ID makes the FamID parameter. If this parameter is not supplied, the stylesheet is applied to all FAM elements in the XML document.
- SortFamilies - accepts either 'true' or 'false'. If the IDs of the FAM elements are structured in a way that can be sorted in ascending or descending order, this parameter tells the stylesheet to sort these IDs and output the families in that order. It defaults to 'false'.

1.1.2. Look and Feel Parameters

- BorderLineStyle - this parameter allows the user to determine the look of the borders in the tables, rows, and cells. It defaults to 'solid', but could accept the following values: none, hidden, dotted, dashed,

¹ <http://www.w3.org/Style/XSL/>

² <http://www.w3.org/TR/xslt>

³ <http://www.neomantic.com/gedcom55XML/>

⁴ <http://www.w3.org/TR/xsl/>

⁵ http://www.adobe.com/devnet/pdf/pdf_reference.html

⁶ <http://xmlgraphics.apache.org/fop/>

⁷ <http://www.progenealogists.com/>

⁸ <http://www.progenealogists.com/familygroup1>

⁹ <http://www.progenealogists.com/familygroup2.pdf>

double, groove, ridge, inset, and outset. Use these values with caution. The stylesheet has been designed assuming solid borders.

- `BorderLineWidth` - this parameter enables the width of the borders to be set. It defaults to .3mm. Supply this parameter with caution. The stylesheet has been designed assuming the default value.

1.2. Stylesheet Versions and Download Locations

The stylesheet described in this document is version 0.1. It is strictly speaking a beta version.

- Both the stylesheet and this documentation is located at <http://www.neomantic.com/downloads/gedcom55XMLtoXSL-FO/gedcom55XMLtoXSL-FO.tar.gz>.
- For verification purposes, I have signed this tar, gzipped archive with my gnupg public key located here¹⁰. The signature of `gedcom55XMLtoXSL-FO.tar.gz` is located at <http://www.neomantic.com/downloads/gedcom55XMLtoXSL-FO/gedcom55XMLtoXSL-FO.tar.gz.sign>.

1.3. Stylesheet License

The source code for is released under the GNU General Public License Version 2¹¹ (GPL). The full text of this license can be found in a file called “`gpl-2.0`” in `gedcom55XMLtoXSL-FO.tar.gz`.

1.4. Stylesheet Updates

Hyperlinks to the most up-to-date version of the stylesheet will be posted to <http://www.neomantic.com>.

2. Usage Instructions

The process of converting a GEDCOM 5.5 file into a pdf document which resembles the progenealogs.com forms takes three steps.

1. Convert the GEDCOM 5.5 file to a GEDCOM 5.5 XML document.
2. Transform the GEDCOM 5.5 XML document to an XSL-FO stylesheet by applying the `gedcom55XMLtoXSL-FO.xsl` stylesheet to the XML document.
3. Process the XSL-FO stylesheet using `fop`.

The instructions below follow several conventions:

- `family.ged` represents a GEDCOM 5.5 file
- `family.xml` represents the `family.ged` file converted into a GEDCOM 5.5 XML document
- `family.fo` represents a XSL-FO document produced by applying the `gedcom55XMLtoXSL-FO.xsl` XSLT stylesheet
- `family.pdf` represents the pdf file that resembles the progenealogist's forms
- Text sandwiched between brackets [] indicates variables that depend upon your computer's environment

¹¹ <http://www.gnu.org/licenses/old-licenses/gpl-2.0.html>

2.1. Convert the GEDCOM 5.5 File to a GEDCOM 5.5 XML Document

To perform the conversion, use a Java application released by Michael H. Kay that converts GEDCOM 5.5 files into what he calls a “GedML”¹², XML documents. GedML is similar to GEDCOM 5.5 XML. See this link¹³ for details.

1. Download Kay's source code¹⁴ and unzip it in a location of your choosing. Remember the path to this location. It will be referred to in the last step using the variable `[path-to-gedml-classes]`.
2. Find the files called “GedcomParser.java” and “GedcomToXml.xsl” in Kay's source code.
3. Compile the file `GedcomParser.java` using your favorite Java distribution's compiler - `javac`. This will produce a class file called “GedcomParser.class”. The command is as follows:

```
javac GedcomParser.java
```

4. Download Kay's `saxon`¹⁵ parser, install it, and remember its location. Its location will be referred to below using the `[path-to-saxon]` variable. (It may already be installed on your system; in the Debian GNU/Linux distribution it is located at `/usr/share/java/saxon.jar`.)
5. Convert `family.ged` to `family.xml` by issuing the following command in your terminal:

```
java -cp [path-to-saxon]/saxon.jar:[path-to-gedml-classes]  
com.icl.saxon.StyleSheet -x GedcomParser -o family.xml family.ged  
[path-to-gedml-classes]/GedcomToXml.xsl
```

The output of this command, `family.xml`, will be a near perfect reproduction of a GEDCOM 5.5 into GEDCOM 5.5 XML.

2.2. Transform the GEDCOM 5.5 XML Document to an XSL-FO stylesheet

To perform the transformation, the `gedcom55XMLtoXSL-FO.xsl` stylesheet must be applied to `family.xml` using an XSLT processor. I have used two command line applications to perform this transformation: `xsltproc`¹⁶ and `saxon`. `xsltproc` uses the Open Source XML parser called “libxml”¹⁷ and is included in many GNU/Linux distributions. SAXON is the XSLT processor developed by Michael H. Kay, used in the previous step.

The `gedcom55XMLtoXSL-FO.xsl` stylesheet parameters described above can be supplied to both XSLT processors. Please consult their respective documentation to find out how to do so.

NOTE: Both `xsltproc` and `saxon` may fail to transform extremely large GEDCOM 5.5 XML files due to either your computer's or the Java virtual machine's memory limitations. To get around these limitations, the `saxon` processor works best because java's memory usage can be manipulated. See the instructions below.

¹² <http://homepage.ntlworld.com/michael.h.kay/gedml/>

¹³ <http://www.neomantic.com/gedcom55XML/0.1/README.html#id2515344>

¹⁶ <http://xmlsoft.org/XSLT/xsltproc2.html>

¹⁷ <http://xmlsoft.org/>

2.2.1. xsltproc

To use `xsltproc`, issue the following command in a terminal:

```
jxsltproc gedcom55toXSL-FO.xsl family.xml > family.fo
```

2.2.2. SAXON

To use `saxon`, issue the following command in a terminal:

```
java -cp [path-to-saxon]/saxon.jar com.icl.saxon.StyleSheet -o  
family.fo family.xml gedcom55toXSL-FO.xsl
```

As mentioned above, if the `family.xml` file is extremely large, this command may fail due to Java's default memory settings. To increase the memory used by the Java's virtual machine, simply add these two flags to the command line: `-Xms -Xmx` and adjust the memory by supplying a value like 512m for each flag.

2.3. Process the XSL-FO Stylesheet Using fop

The outcome of the previous step is an XSL-FO document called `family.fo`. This file now needs to be processed by `fop` to produce the pdf document. To do so, follow these steps:

1. Follow the `fop` “quickstart” guide here¹⁸ to download and install `fop` or use your GNU/Linux distribution's package management system to install it.
2. Once installed, issue the following command in a terminal:

```
fop -fo family.fo -pdf family.pdf
```

The output of this command is the `family.pdf` file.

2.4. Sample

Included with the stylesheet and this documentation is a pdf sample of the output. It is based on a GEDCOM 5.5 file called “royal92.ged” which contains this genealogical history of the British royal family. This GEDCOM 5.5 can be easily found on the web.

3. Documentation License

This document is released under the GNU Free Documentation License Version 1.2¹⁹. The full text of this license is found in the file called “`fdl.txt`” released with `gedcom55XMLtoXSL-FO.tar.gz`. It can also be located at <http://www.neomantic.com/downloads/gedcom55XMLtoXSL-FO/README.html>²⁰.

4. Contact

Please direct questions or requests for more information to [<chad@neomantic.com>](mailto:chad@neomantic.com). Corrections, suggestions, bug reports, and patches are welcome as well.

¹⁹ <http://www.gnu.org/licenses/fdl.html>

²⁰ <http://www.neomantic.com/downloads/gedcom55XMLtoXSL-FO/README.html>