

Creating native Excel workbooks on a non-Windows system without SAS/ACCESS



Edwin van Stein
Astellas Pharma
PhUSE 2015-10-14

© October 2015 Edwin van Stein

The views reflected here are personal and are not those of my employer, Astellas Pharma Europe B.V.

Where copyright subsists in any of the material shown here today original copyright and/or ownership is hereby acknowledged.



Customer wants

- All SDTM data sets for a study
- ... in one file
- ... to be opened in Excel
- ... with auto filter turned on
- ... with proper column widths
- ... not compressed



XLS
SAS/ACCESS

CSV
DATA step
PROC EXPORT
ODS CSVALL

XML
TAGSETS.EXCELXP
TAGSETS.EXCELBASE



Office Open XML

ISO/IEC 29500 standard

Freely available

Includes DOCX, PPTX, XLSX

*It's just a ZIP file
containing XML files!*



Creating XLSX in SAS

Create directory structure (X commands)



Create XML files (DATA steps)



Combine with zip into XLSX file (X commands)

XLSX file with 3 worksheets

/_rels/.rels

/docProps/app.xml

/docProps/core.xml

/xl/_rels/workbook.xml.rels

/xl/worksheets/sheet1.xml

/xl/worksheets/sheet2.xml

/xl/worksheets/sheet3.xml

/xl/workbook.xml

/[Content_types].xml

The easy parts

No modifications needed

/_rels/.rels

Modifications possible, but not necessary

/docProps/app.xml

/docProps/core.xml

Only variable part: number of sheets

/[Content_types].xml

/xl/_rels/workbook.xml.rels

Another easy part

/xl/workbook.xml contains:

- worksheet names (max 31 characters)

- worksheet IDs

The most interesting parts

/xl/worksheets/worksheet1.xml contains:

- frozen panes

- auto filters

- column widths

- the actual data!

10 slides, still no code?!



	A	B	C	D	E
1	Name	Sex	Age	Height	Weight
2	Name 	Sex 	Age 	Height 	Weight 
3	Alfred	M	14	69	112.5
4	Alice	F	13	56.5	84
5	Barbara	F	13	65.3	98
6	Carol	F	14	62.8	102.5
7	Henry	M	14	63.5	102.5
8	James	M	12	57.3	83
9	Jane	F	12	59.8	84.5
10	Janet	F	15	62.5	112.5
11	Jeffrey	M	13	62.5	84
12	John	M	12	59	99.5
13	Joyce	F	11	51.3	50.5
14	Judy	F	14	64.3	90
15	Louise	F	12	56.3	77
16	Mary	F	15	66.5	112
17	Philip	M	16	72	150
18	Robert	M	12	64.8	128
19	Ronald	M	15	67	133
20	Thomas	M	11	57.5	85
21	William	M	15	66.5	112

<?xml>

<worksheet>

<dimension />

← dimension of the worksheet

<sheetviews>

<sheetview>

<pane />

← frozen pane information

<selection />

← selected cell(s) information

</sheetview>

</sheetviews>

<cols>

<col />

← column width information

</cols>

<sheetData>

<row>

<c />

← actual cell values

</row>

</sheetData>

<autofilter />

← auto filter information

</worksheet>

| | A | B | C | D | E |
|---|--------|-------|-------|----------|----------|
| 1 | Name | Sex | Age | Height | Weight |
| 2 | Name ▼ | Sex ▼ | Age ▼ | Height ▼ | Weight ▼ |
| 3 | Alfred | M | 14 | 69 | 112.5 |
| 4 | Alice | F | 13 | 56.5 | 84 |

```

<dimension ref="A1:E21"/>
<sheetViews>
<sheetView tabSelected="1"
workbookViewId="0">
<pane ySplit="2" topLeftCell="A3"
activePane="bottomLeft" state="frozen"/>
<selection pane="bottomLeft"
activeCell="A3" sqref="A3"/>
</sheetView>
</sheetViews>
<cols>
<col min="1" max="1" width="13"
customWidth="1"/>

```

| | A | B | C | D | E |
|---|--------|-----|-----|--------|--------|
| 1 | Name | Sex | Age | Height | Weight |
| 2 | Name | Sex | Age | Height | Weight |
| 3 | Alfred | M | 14 | 69 | 112.5 |
| 4 | Alice | F | 13 | 56.5 | 84 |

```
<sheetData>
```

```
<row r="3">
```

```
<c r="A3"
```

```
t="inlineStr"><is><t>Alfred</t></is></c>
```

```
<c r="B3" t="inlineStr"><is><t>M</t></is></c>
```

```
<c r="C3"><v>14</v></c>
```

```
<c r="D3"><v>69</v></c>
```

```
<c r="E3"><v>112.5</v></c>
```

```
</row>
```

```
</sheetData>
```

| | A | B | C | D | E |
|---|--------|-------|-------|----------|----------|
| 1 | Name | Sex | Age | Height | Weight |
| 2 | Name ▼ | Sex ▼ | Age ▼ | Height ▼ | Weight ▼ |
| 3 | Alfred | M | 14 | 69 | 112.5 |
| 4 | Alice | F | 13 | 56.5 | 84 |

```
<autoFilter ref="A2:E21"/>
```

```
</worksheet>
```

Creating the XML files in SAS

```
data _null_;  
    file "&dir./docProps/core.xml"  
        encoding="utf-8" lrecl=32000;  
    put '<?xml version="1.0" ' @;  
    put 'encoding="UTF-8" ' @;  
    put 'standalone="yes"?>' ' 0D'x;  
run;
```

SAS does not like [Content_types].xml

Creating the XML files in SAS

Encoding characters:

```
htmlencode(variable, 'amp gt lt apos quot 7bit')
```

Minimizing redundant spaces:

```
put variable +(-1);
```

Sheet name rules:

```
compress(htmlencode(substr(variable,  
1,31), 'amp gt lt apos quot 7bit'), '/\?  
*[]:')
```

Creating the XML files in SAS

Encoding issues:

```
if index(variable, '1A'x) then  
variable=translate(variable, '?', '1A'x);
```

Converting numbers to Excel columns:

byte(65) is the letter A, byte(90) is the letter Z

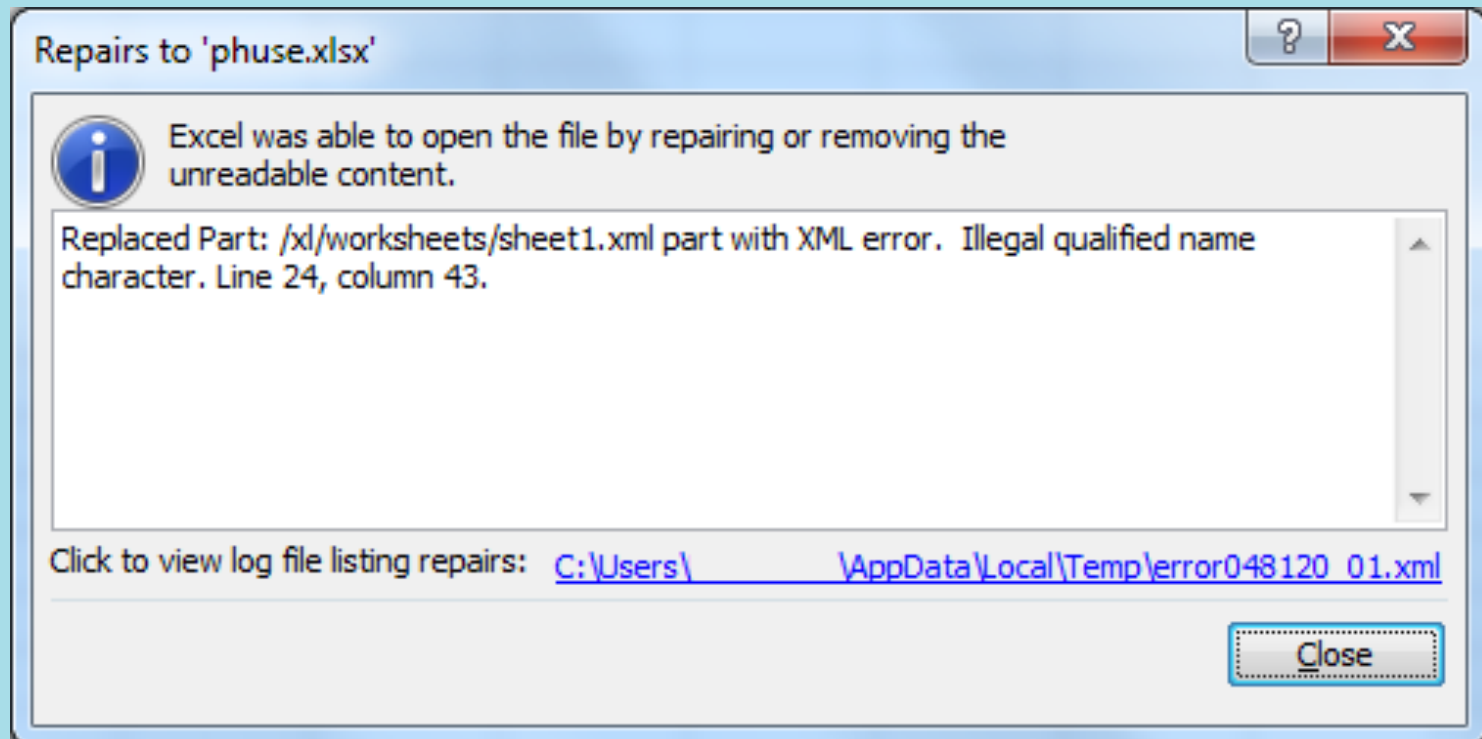
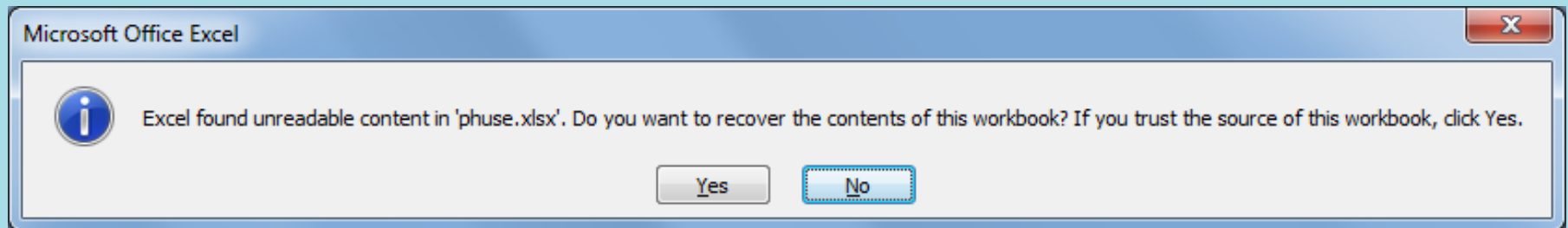
```
if col le 26 then __col=byte(col+64);  
else __col=translate(byte(floor((col-1)/  
26)+64) !! byte(mod(col,26)+64), 'Z', '@');
```

Finishing up

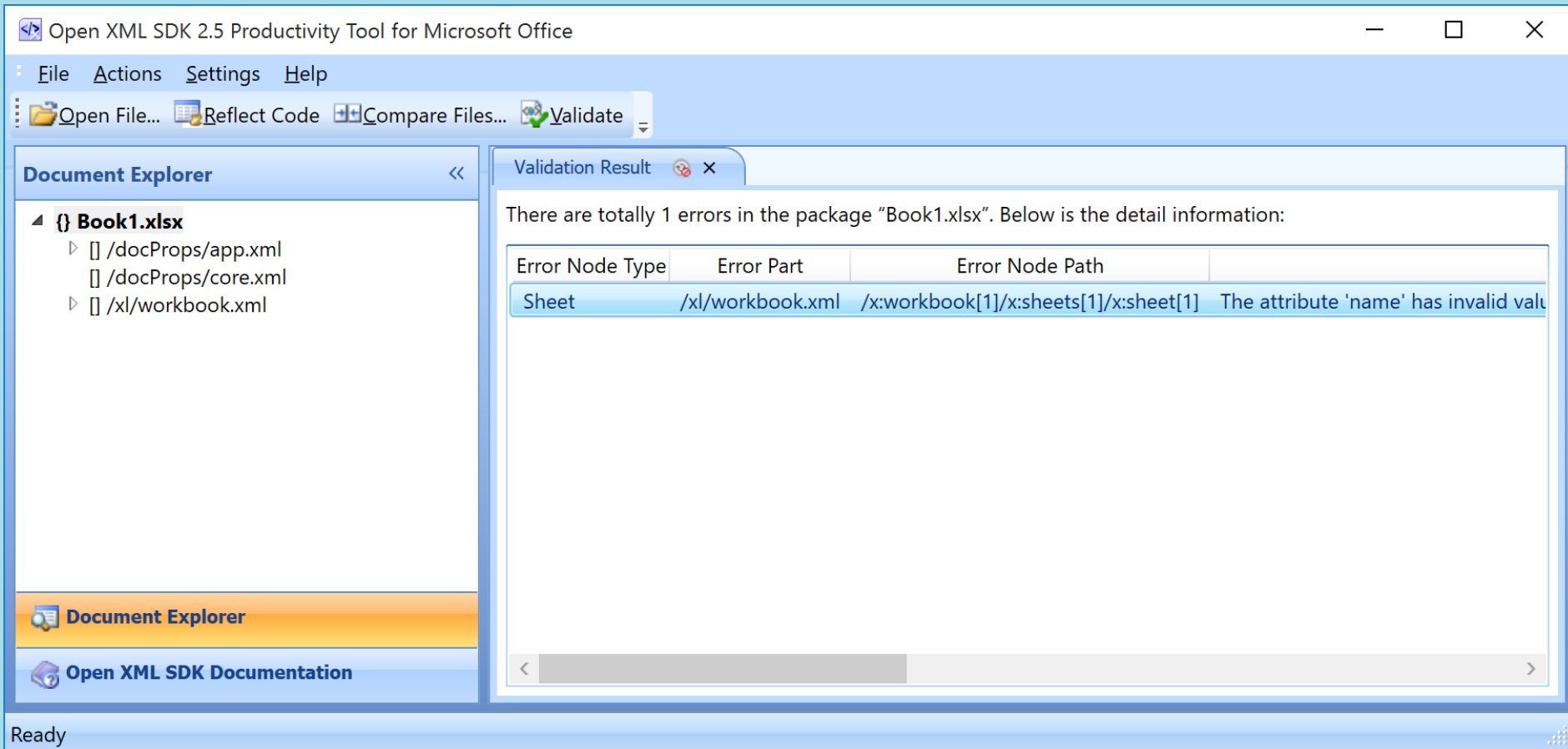
Zip the XML files including directory structure

```
x "cd '&dir.'; zip -r &outfile..xlsx *";
```

Debugging with Excel



Open XML SDK 2.5 Productivity Tool for Microsoft Office





Edwin van Stein

edwin-vanstein (at) astellas.com

<http://nl.linkedin.com/in/ejvanstein>

images

designed by freepik.com

