Easy Hacks to Improve Writer - OOXML Interoperability

LibreOffice Conference 2014, Bern

Sushil Shinde
sushil.shinde@synerzip.com

SYNERZIP

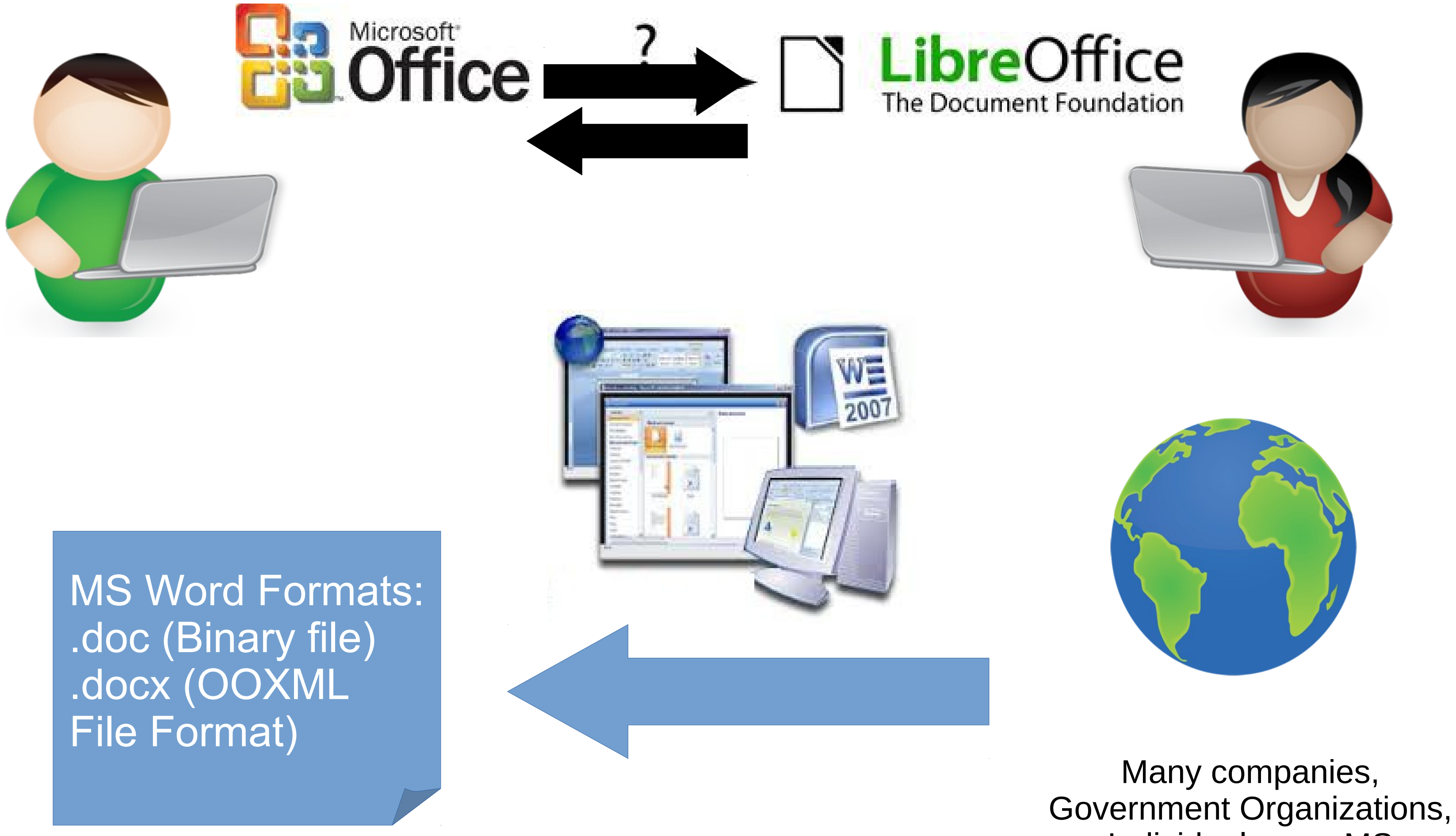
About Me

- Software Developer at Synerzip Softech India
- About 3 years of experience in C++ and OOXML
- Active contributor to LibreOffice product and community
- Member of TDF.
- Love to play, watch cricket
- Email: Sushil.shinde@synerzip.com
- IRC: #libreoffice-dev chat : sushils_

Topics

- Interoperability
- OOXML and ECMA-376
- DOCX File Structure
- Challenges during 'File Import'
  - File Crash
  - Data Loss
- Challenges during 'File Export'
  - File Corruption
  - Data Loss
- LibreOffice Hang Issues
- Some Useful Tools
- Examples

Interoperability



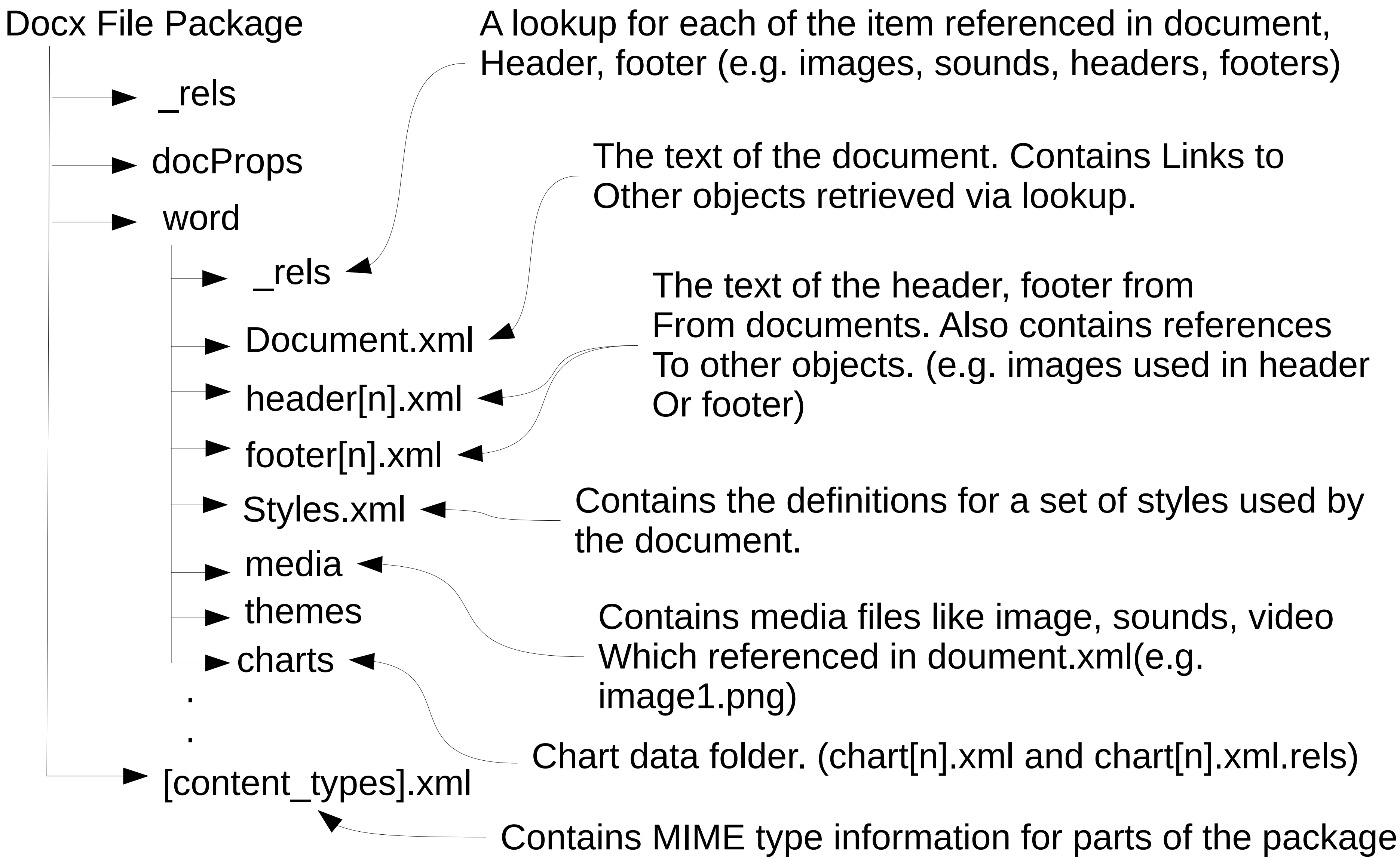
MS Word Formats:
.doc (Binary file)
.docx (OOXML
File Format)

Many companies,
Government Organizations,
Individuals use MS
Word File Formats.

OOXML and ECMA-376

- Office Open XML (OOXML)
  - Microsoft Office 2007 and later versions (like 2010, 2013) uses OOXML format.
- The ECMA-376 Standard
  - This Standard defines OOXML's vocabularies and document representation and packaging details.
  - Specifications are freely available on the ECMA website.

DOCX File Structure



Docx File Package

- _rels
- docProps
- word
  - _rels
  - Document.xml
  - header[n].xml
  - footer[n].xml
  - Styles.xml
  - media
  - themes
  - charts
  - .
  - .
- [content_types].xml

A lookup for each of the item referenced in document, Header, footer (e.g. images, sounds, headers, footers)

The text of the document. Contains Links to Other objects retrieved via lookup.

The text of the header, footer from From documents. Also contains references To other objects. (e.g. images used in header Or footer)

Contains the definitions for a set of styles used by the document.

Contains media files like image, sounds, video Which referenced in doument.xml(e.g. image1.png)

Chart data folder. (chart[n].xml and chart[n].xml.rels)

Contains MIME type information for parts of the package

Challenges In 'File Import'

- LibreOffice crash
- Data loss
- LibreOffice hangs

File Import – Crash issues

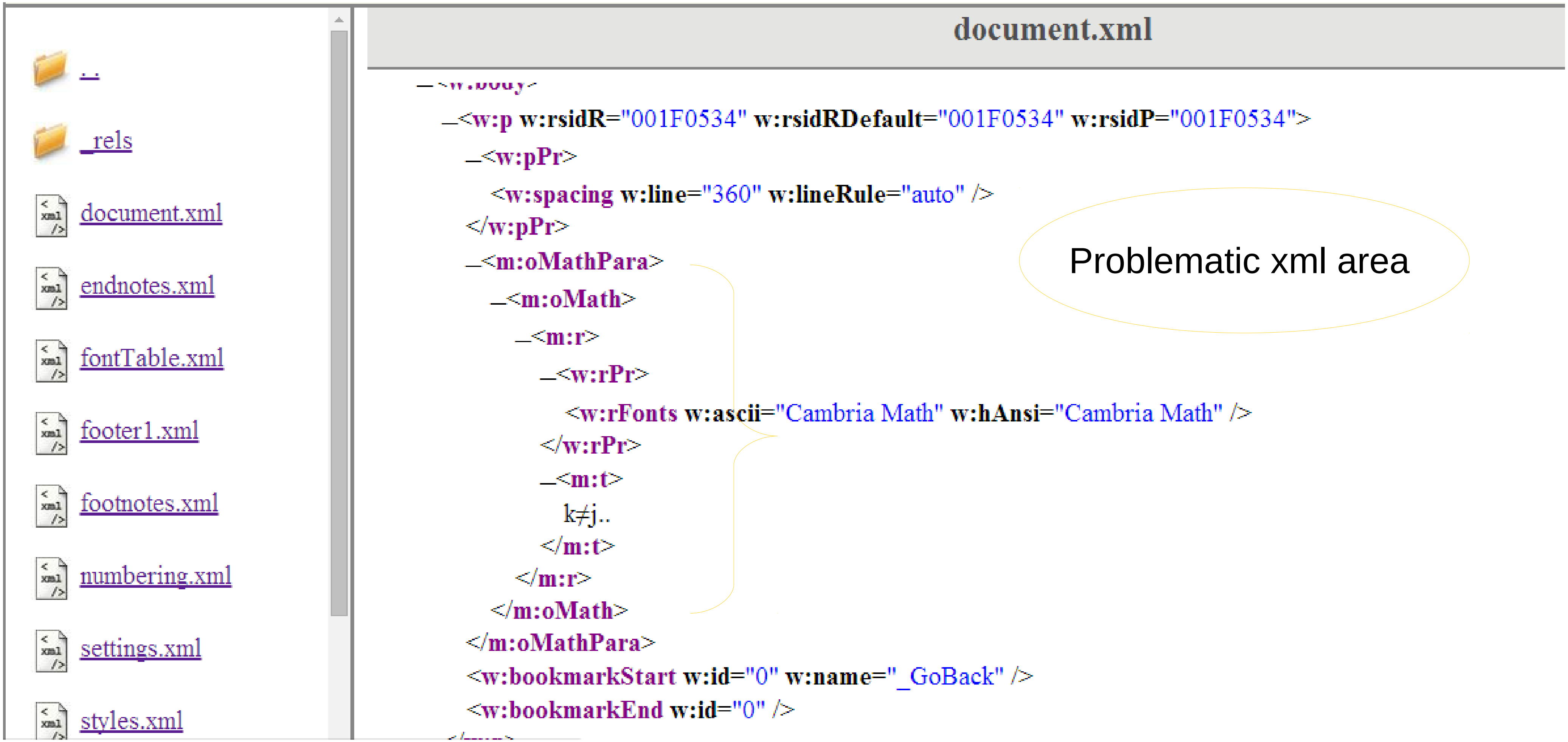
- Reasons can be-
  - Programming mistakes
    - Null pointer check
    - Memory Leaks
  - Some issues in import filters
    - Some specific combinations of data

Analyzing Crash

- Optimize File
  - Check MS Office version (2007/2010/2013) using which file is created
  - Use “Divide and conquer” method to optimize file
  - Try to optimize file upto 1-2 pages with minimum data on it
- Identify XML part which is causing error
- Try to Identify MS Office feature which is causing error
  - If confirmed, try to create .doc (binary version) file with same feature and check whether that file works
- Locate parsing and mapping of XML elements in import filters to identify root cause

Crash - Example

fdo#79973



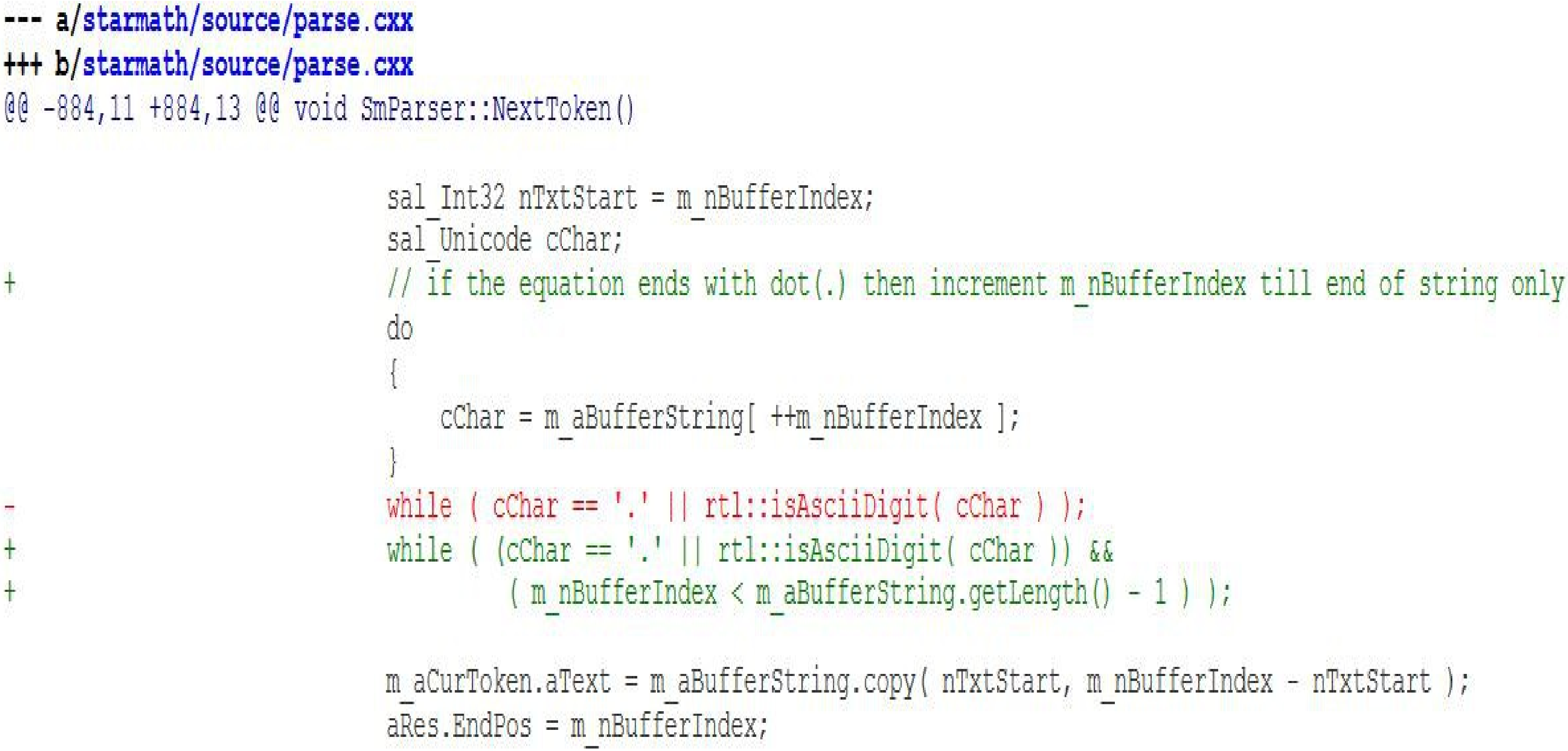
- ..
- _rels
- document.xml
- endnotes.xml
- fontTable.xml
- footer1.xml
- footnotes.xml
- numbering.xml
- settings.xml
- styles.xml

document.xml

```
– <w:body>
  – <w:p w:rsidR="001F0534" w:rsidRDefault="001F0534" w:rsidP="001F0534">
    – <w:pPr>
        <w:spacing w:line="360" w:lineRule="auto" />
      </w:pPr>
    – <m:oMathPara>
      – <m:oMath>
        – <m:r>
          – <w:rPr>
              <w:rFonts w:ascii="Cambria Math" w:hAnsi="Cambria Math" />
            </w:rPr>
          – <m:t>
              k≠j..
            </m:t>
          </m:r>
        </m:oMath>
      </m:oMathPara>
      <w:bookmarkStart w:id="0" w:name="_GoBack" />
      <w:bookmarkEnd w:id="0" />
    </w:p>
```

Problematic xml area

Resolving Crash - Example



```
--- a/starmath/source/parse.cxx
+++ b/starmath/source/parse.cxx
@@ -884,11 +884,13 @@ void SmParser::NextToken()

                sal_Int32 nTxtStart = m_nBufferIndex;
                sal_Unicode cChar;
+               // if the equation ends with dot(.) then increment m_nBufferIndex till end of string only
                do
                {
                    cChar = m_aBufferString[ ++m_nBufferIndex ];
                }
-               while ( cChar == '.' || rtl::isAsciiDigit( cChar ) );
+               while ( (cChar == '.' || rtl::isAsciiDigit( cChar )) &&
+                        ( m_nBufferIndex < m_aBufferString.getLength() - 1 ) );

                m_aCurToken.aText = m_aBufferString.copy( nTxtStart, m_nBufferIndex - nTxtStart );
                aRes.EndPos = m_nBufferIndex;
```

Code reference : https://gerrit.libreoffice.org/#/c/9840

File Import – Types Of Data Loss

- Feature loss (ex. Text, shapes etc)
- Feature property loss (ex. Colors, line styles etc)
- Incorrect values (ex. Shape size, position etc)

File Import – Reasons For Data Loss

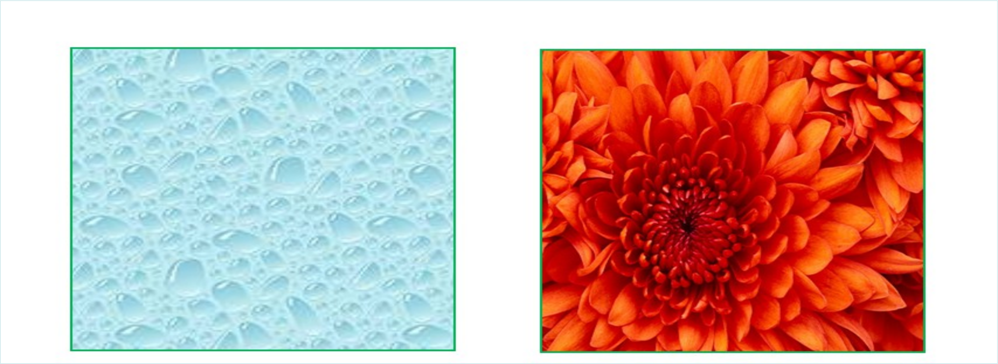
- MS Office feature is not supported
  - Implement feature support
  - Grab-bag
- XML Nodes not handled
- XML elements not mapped properly
- Properties lost in shape conversions (SwXShape → SwXTextFrame)

File Import – How To Fix Data Loss

- Check XML Schema of missing feature
- Check ECMA 376 specs of missing properties
- Check XML properties are available in model.xml
- Identify LibreOffice UNO Properties for missing data
  - Insert similar feature in LibreOffice and check properties that represent missing effects
  - Create .doc file with same data
  - Use XRAY tool to check properties
- Locate handling of those XML properties in dmapper
- Check XML values are properly mapped with UNO properties
  - Hard-code UNO Properties to verify quickly

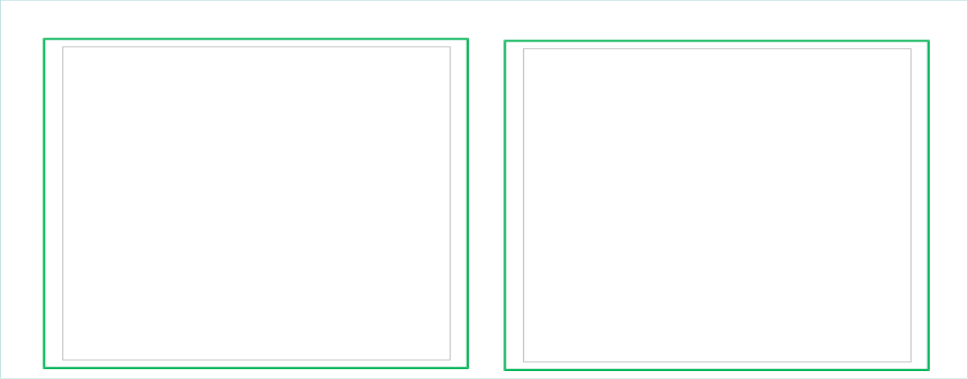
Data Loss Example - shape

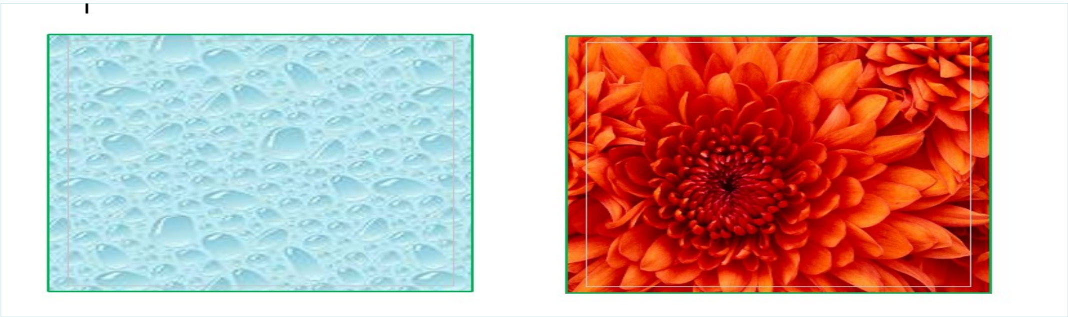
- TextBox Background image loss



Original TextBox fill

LO rendered before FIX





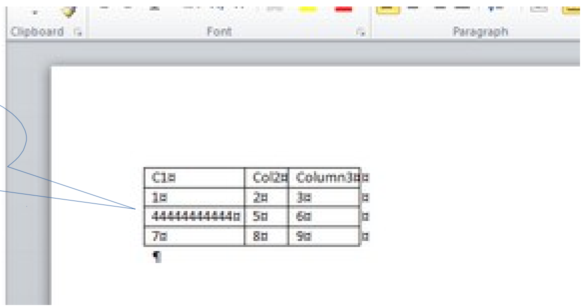
LO rendered after fix

Data Loss Example - shape

- Set proper UNO Property
  - “FillBitmapURL” property for shape
  - “BackGraphicURL” property for TextFrame
- Handled “BackGraphicURL” property in export if it is textframe

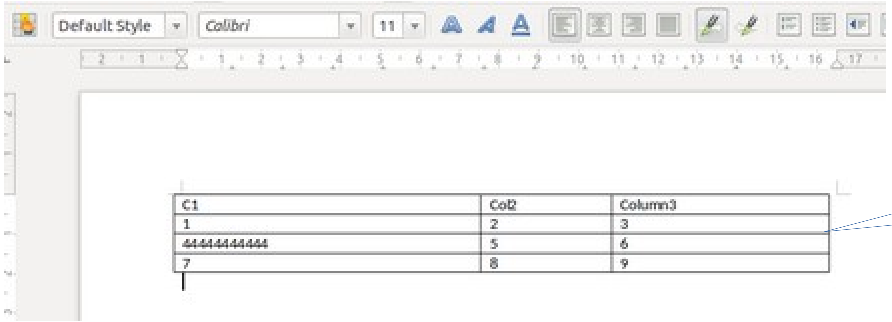
  Code Reference : https://gerrit.libreoffice.org/#/c/7259

Data Loss Example - Table



C1¤	Col2¤	Column3¤¤
1¤	2¤	3¤
44444444444¤	5¤	6¤
7¤	8¤	9¤

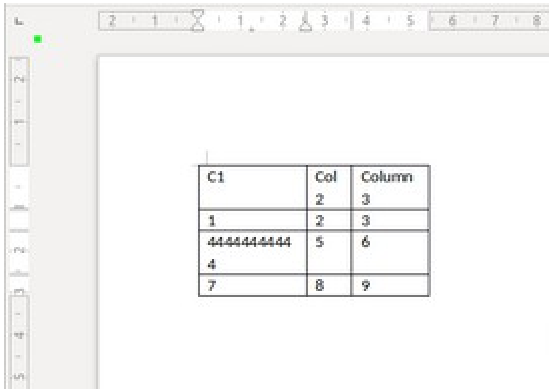
Original table
Auto width



C1	Col2	Column3
1	2	3
44444444444	5	6
7	8	9

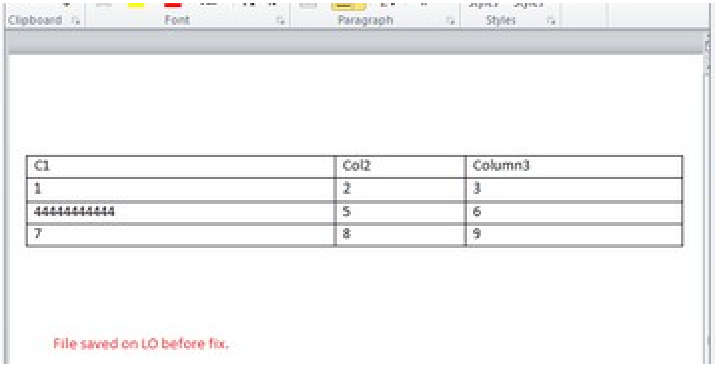
How LO rendered

LO Rendering After Fix



C1	Col2	Column3
1	2	3
44444444444	5	6
7	8	9

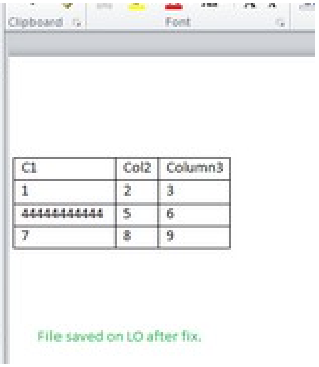
LO : Export Before Fix



C1	Col2	Column3
1	2	3
44444444444	5	6
7	8	9

File saved on LO before fix.

After Fix



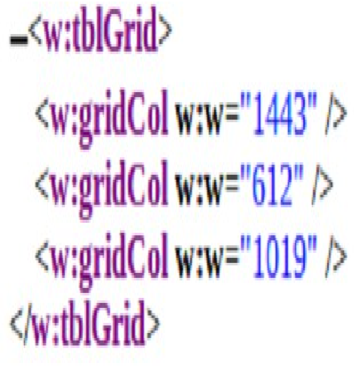
C1	Col2	Column3
1	2	3
44444444444	5	6
7	8	9

File saved on LO after fix.

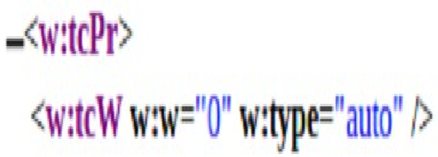
Data Loss Example - Table

XML Comparison

Original

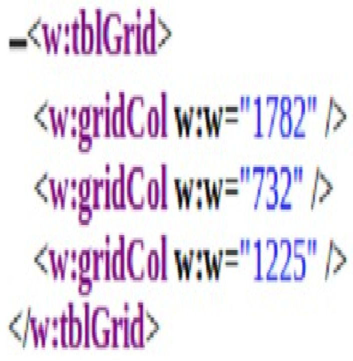


```
-<w:tblGrid>
  <w:gridCol w:w="1443" />
  <w:gridCol w:w="612" />
  <w:gridCol w:w="1019" />
</w:tblGrid>
```

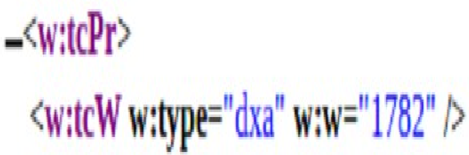


```
-<w:tcPr>
  <w:tcW w:w="0" w:type="auto" />
```

LO Exported this..

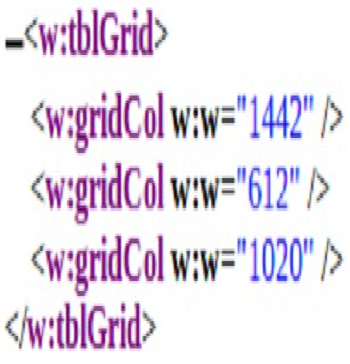


```
-<w:tblGrid>
  <w:gridCol w:w="1782" />
  <w:gridCol w:w="732" />
  <w:gridCol w:w="1225" />
</w:tblGrid>
```

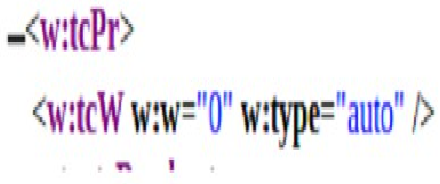


```
-<w:tcPr>
  <w:tcW w:type="dxa" w:w="1782" />
```

Fixed



```
-<w:tblGrid>
  <w:gridCol w:w="1442" />
  <w:gridCol w:w="612" />
  <w:gridCol w:w="1020" />
</w:tblGrid>
```



```
-<w:tcPr>
  <w:tcW w:w="0" w:type="auto" />
```

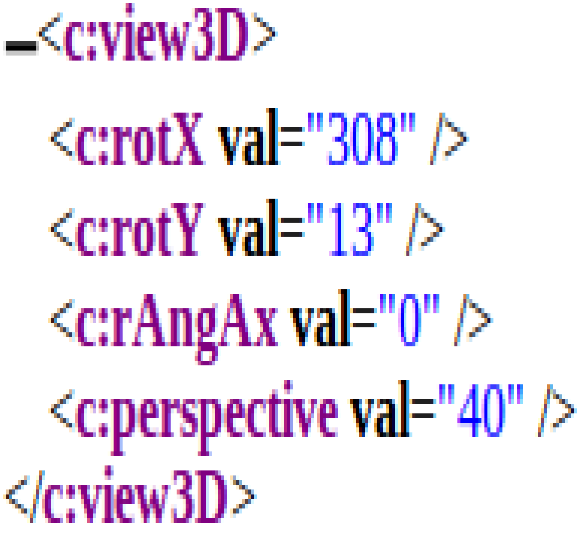
Code Reference : https://gerrit.libreoffice.org/#/c/7593/
https://gerrit.libreoffice.org/#/c/7594/

Challenges In 'File Export'

- MS Office not able to open 'saved file'
- Data loss
- LO crash

File Export – Types Of Corruptions

- Invalid XML values exported
  - XML values are not exported as per ECMA specs

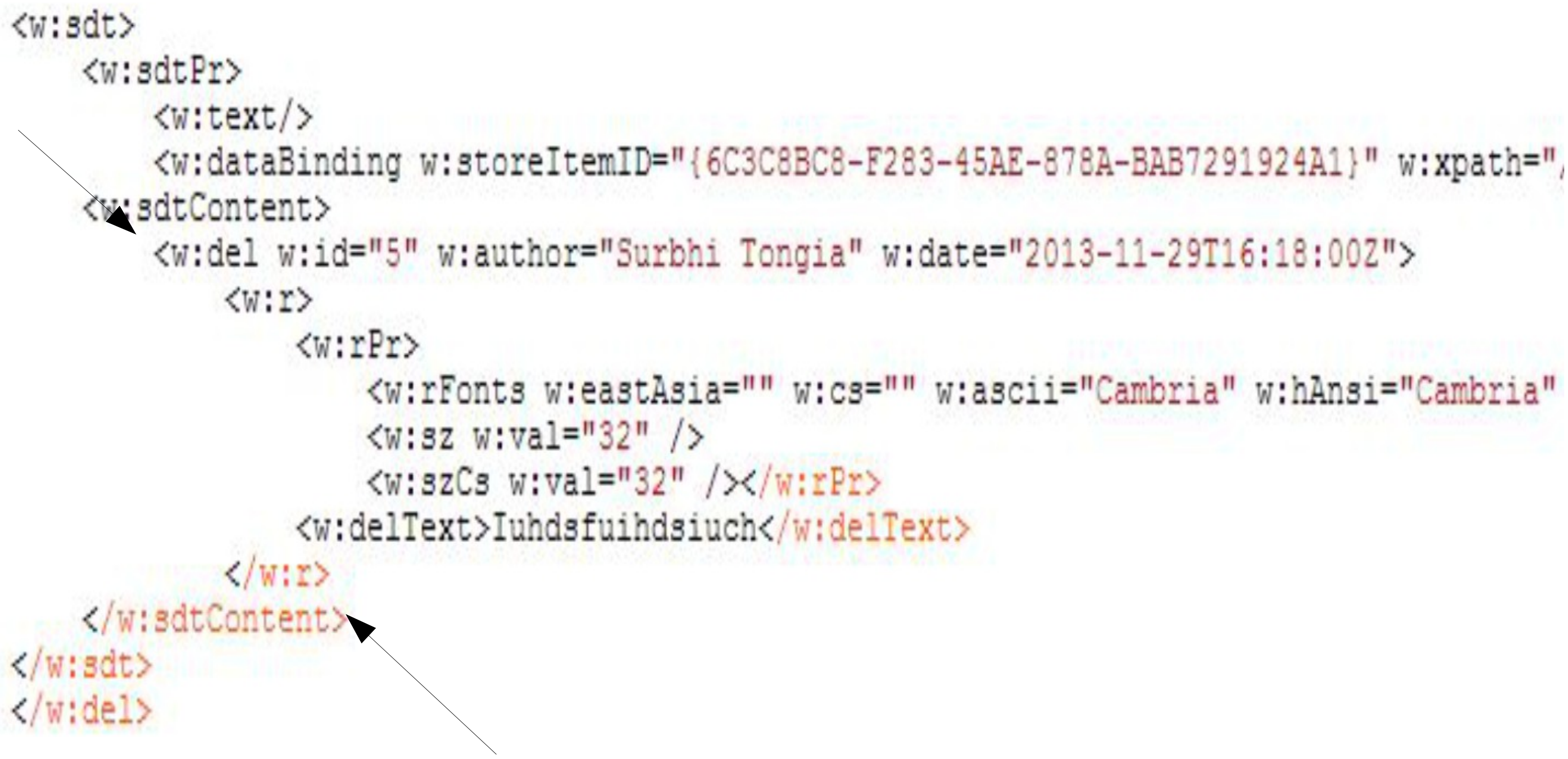


```
- <c:view3D>
    <c:rotX val="308" />
    <c:rotY val="13" />
    <c:rAngAx val="0" />
    <c:perspective val="40" />
  </c:view3D>
```

ECMA specs : valid values for rotX are between [-90,90]

File Export – Types Of Corruptions

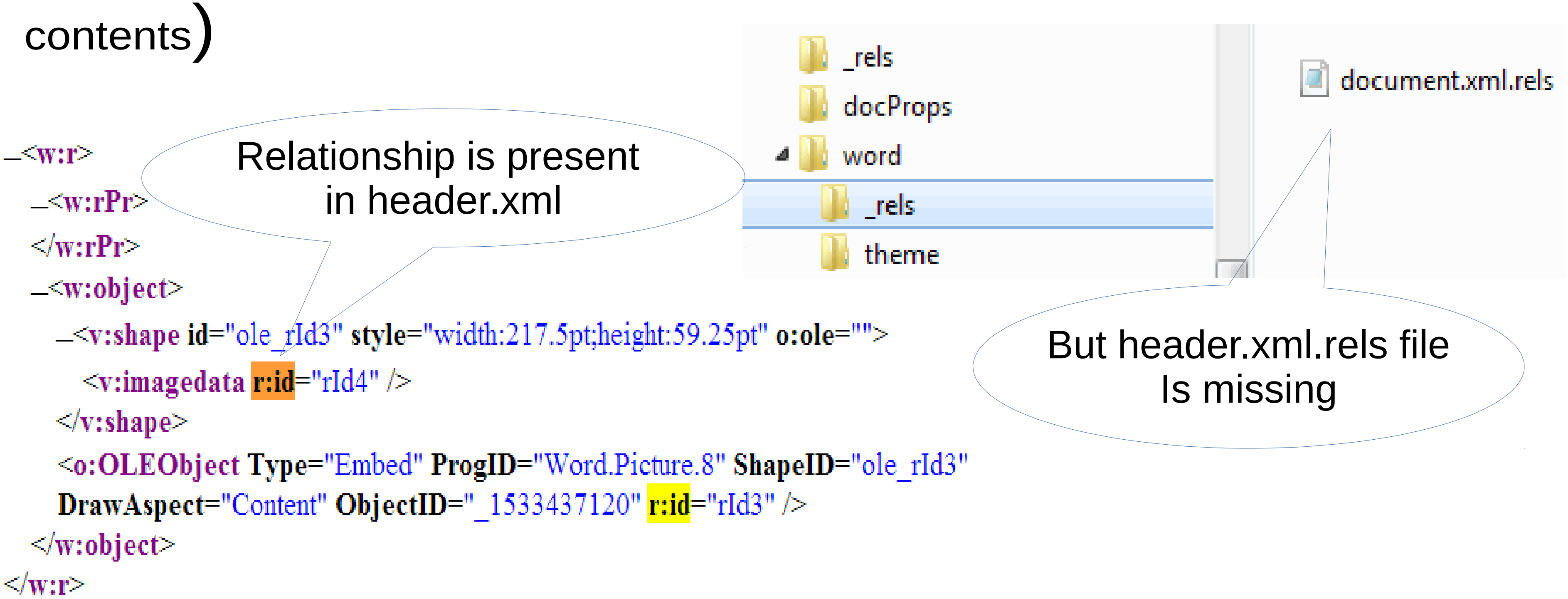
- XML tag mismatch – Start and End tag not matching



```
<w:sdt>
    <w:sdtPr>
        <w:text/>
        <w:dataBinding w:storeItemID="{6C3C8BC8-F283-45AE-878A-BAB7291924A1}" w:xpath="
    <w:sdtContent>
        <w:del w:id="5" w:author="Surbhi Tongia" w:date="2013-11-29T16:18:00Z">
            <w:r>
                <w:rPr>
                    <w:rFonts w:eastAsia="" w:cs="" w:ascii="Cambria" w:hAnsi="Cambria"
                    <w:sz w:val="32" />
                    <w:szCs w:val="32" /></w:rPr>
                <w:delText>Iuhdsfuihdsiuch</w:delText>
            </w:r>
    </w:sdtContent>
</w:sdt>
</w:del>
```

File Export – Types Of Corruptions

- Missing target relationship entry
- Missing relationship file (ex. header.xml.rels)
- Exported 0 bytes file (Mostly in case of images/media folder contents)



```
–<w:r>
  –<w:rPr>
  </w:rPr>
  –<w:object>
    –<v:shape id="ole_rId3" style="width:217.5pt;height:59.25pt" o:ole="">
       <v:imagedata r:id="rId4" />
    </v:shape>
    <o:OLEObject Type="Embed" ProgID="Word.Picture.8" ShapeID="ole_rId3"
    DrawAspect="Content" ObjectID="_1533437120" r:id="rId3" />
  </w:object>
</w:r>
```

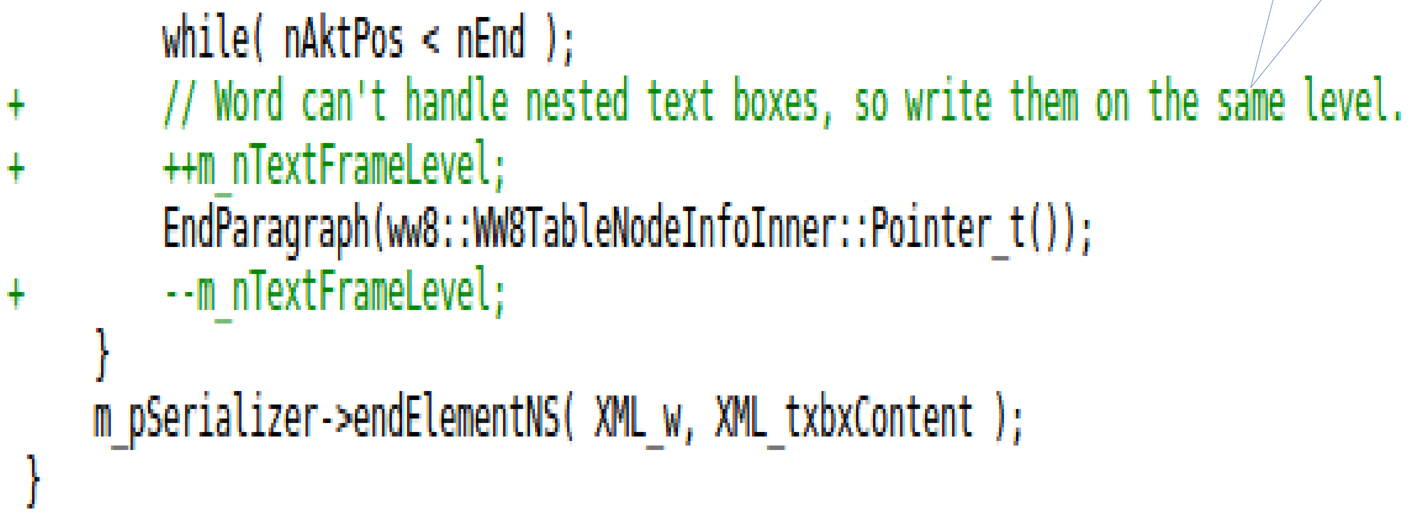
Relationship is present in header.xml

But header.xml.rels file Is missing

File Export – Types Of Corruptions

- Invalid hierarchy
  - Text box exported inside the another textbox

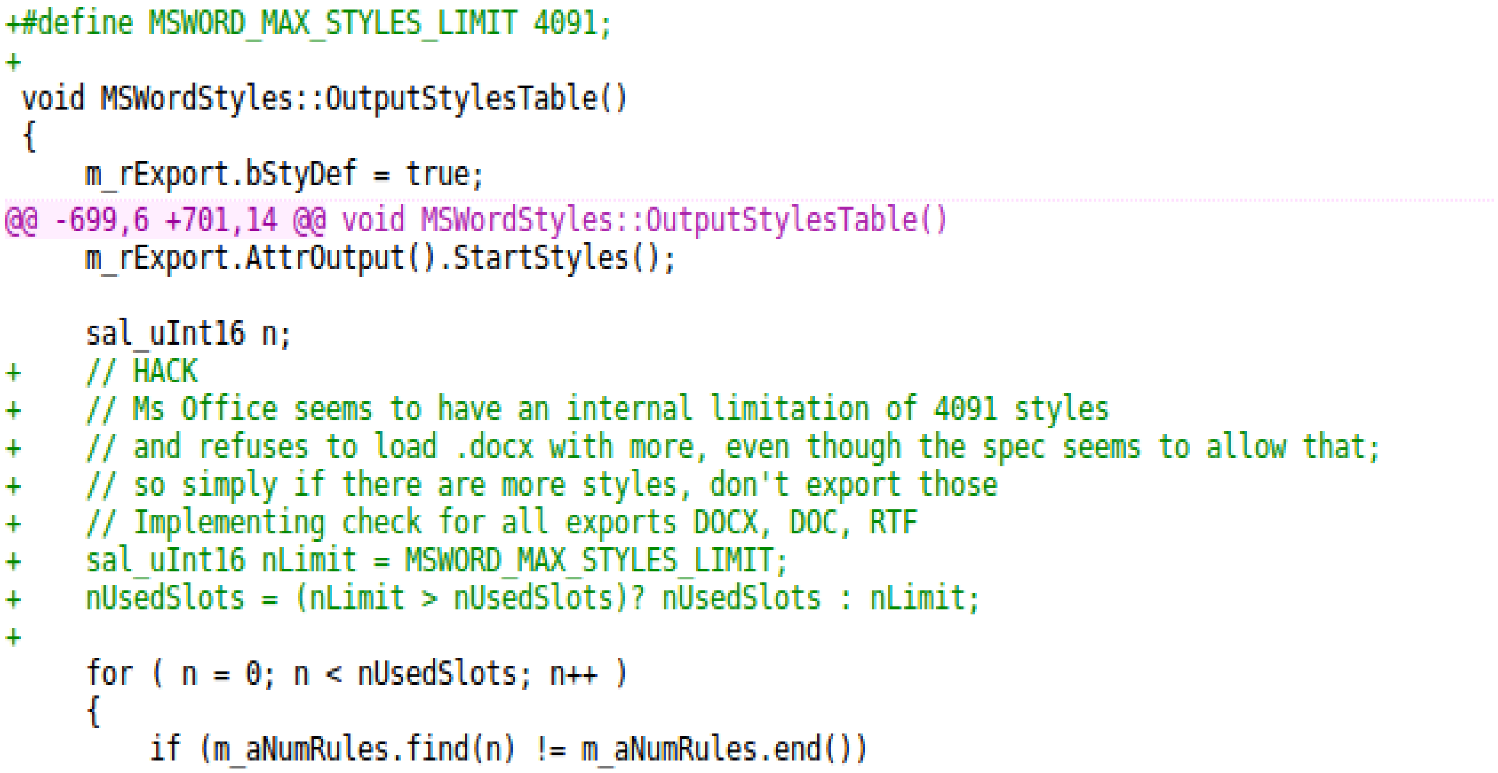
Easy Hack



```
        while( nAktPos < nEnd );
+       // Word can't handle nested text boxes, so write them on the same level.
+       ++m_nTextFrameLevel;
        EndParagraph(ww8::WW8TableNodeInfoInner::Pointer_t());
+       --m_nTextFrameLevel;
    }
    m_pSerializer->endElementNS( XML_w, XML_txbxContent );
 }
```

File Export – Corruption Issues

Ms Office seems to have an internal limitation of 4091 styles and refuses to load ".docx" with more styles.



```
+#define MSWORD_MAX_STYLES_LIMIT 4091;
+
 void MSWordStyles::OutputStylesTable()
 {
     m_rExport.bStyDef = true;
@@ -699,6 +701,14 @@ void MSWordStyles::OutputStylesTable()
     m_rExport.AttrOutput().StartStyles();

     sal_uInt16 n;
+    // HACK
+    // Ms Office seems to have an internal limitation of 4091 styles
+    // and refuses to load .docx with more, even though the spec seems to allow that;
+    // so simply if there are more styles, don't export those
+    // Implementing check for all exports DOCX, DOC, RTF
+    sal_uInt16 nLimit = MSWORD_MAX_STYLES_LIMIT;
+    nUsedSlots = (nLimit > nUsedSlots)? nUsedSlots : nLimit;
+
     for ( n = 0; n < nUsedSlots; n++ )
     {
         if (m_aNumRules.find(n) != m_aNumRules.end())
```

Analyzing File Corruption

- Validate exported docx file
  - Use OpenSDK tool to validate file (For windows only)
- Compare content of exported file with original file
  - Use OOXML tool to compare file
- Check ECMA specs of invalid XML property
- Check relID's are exported properly
  - Relationship target is present in rels xml file
  - Check target file is available in exported file
- Search for export part of invalid XML in export files e.g. docxattributeoutput, docxsdrexport etc.

File Export – Reasons For Data Loss

- Features rendered properly are mostly preserved in export
- Reasons for Data loss can be-
  - Mapping of UNO Properties to OOXML properties
    - Invalid data conversion (from LO property to MSO valid XML value as per ECMA)
    - e.g. Rotation Angle, Dashed Borders etc
  - Required XML part is missing in exported file
    - e.g. Fill properties from shape XML Schema

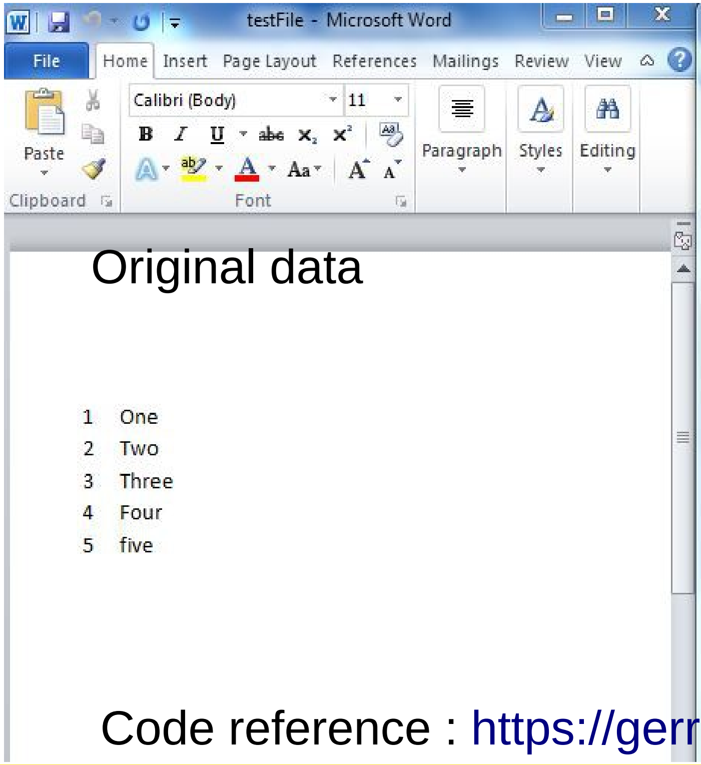
File Export - How To Fix Data Loss

- Compare exported and original file
  - Verify XML schema for missing feature or properties of missing feature are exported
- Check export code for missing XML part.
  - Search for xml tag “XML_elementname” e.g. XML_rot. In export classes.
  - Check xml parts are written under right parent elements.

Data Loss - Example

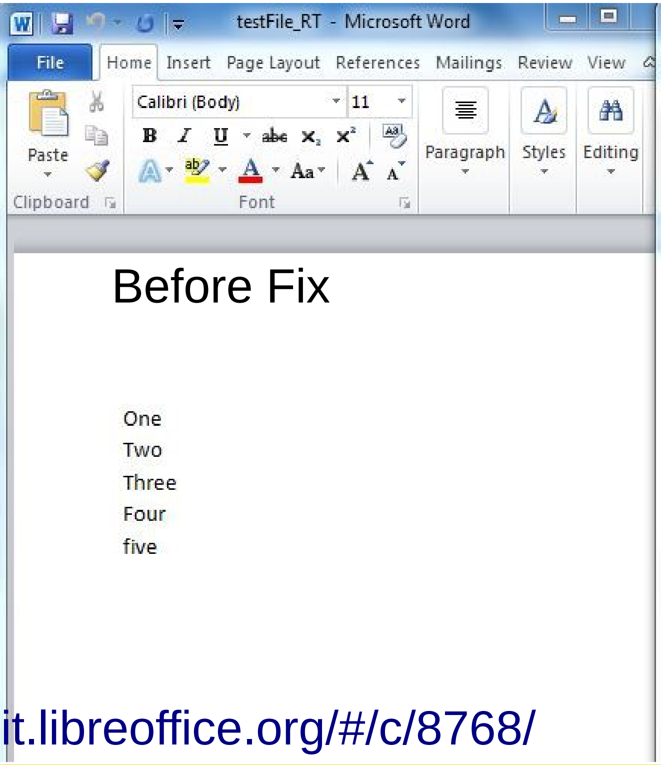
- Numbered list is not preserved
  - Original XML - <w:lvlText w:val="%1" />
  - Exported XML - <w:lvlText w:val="" />

Numbering.xml



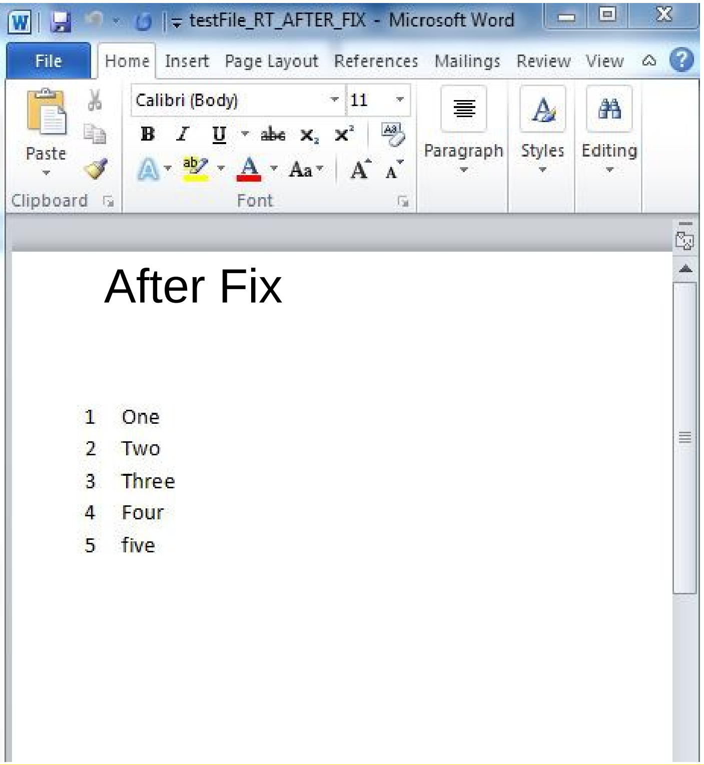
Original data

1 One
2 Two
3 Three
4 Four
5 five



Before Fix

One
Two
Three
Four
five



After Fix

1 One
2 Two
3 Three
4 Four
5 five

Code reference : https://gerrit.libreoffice.org/#/c/8768/

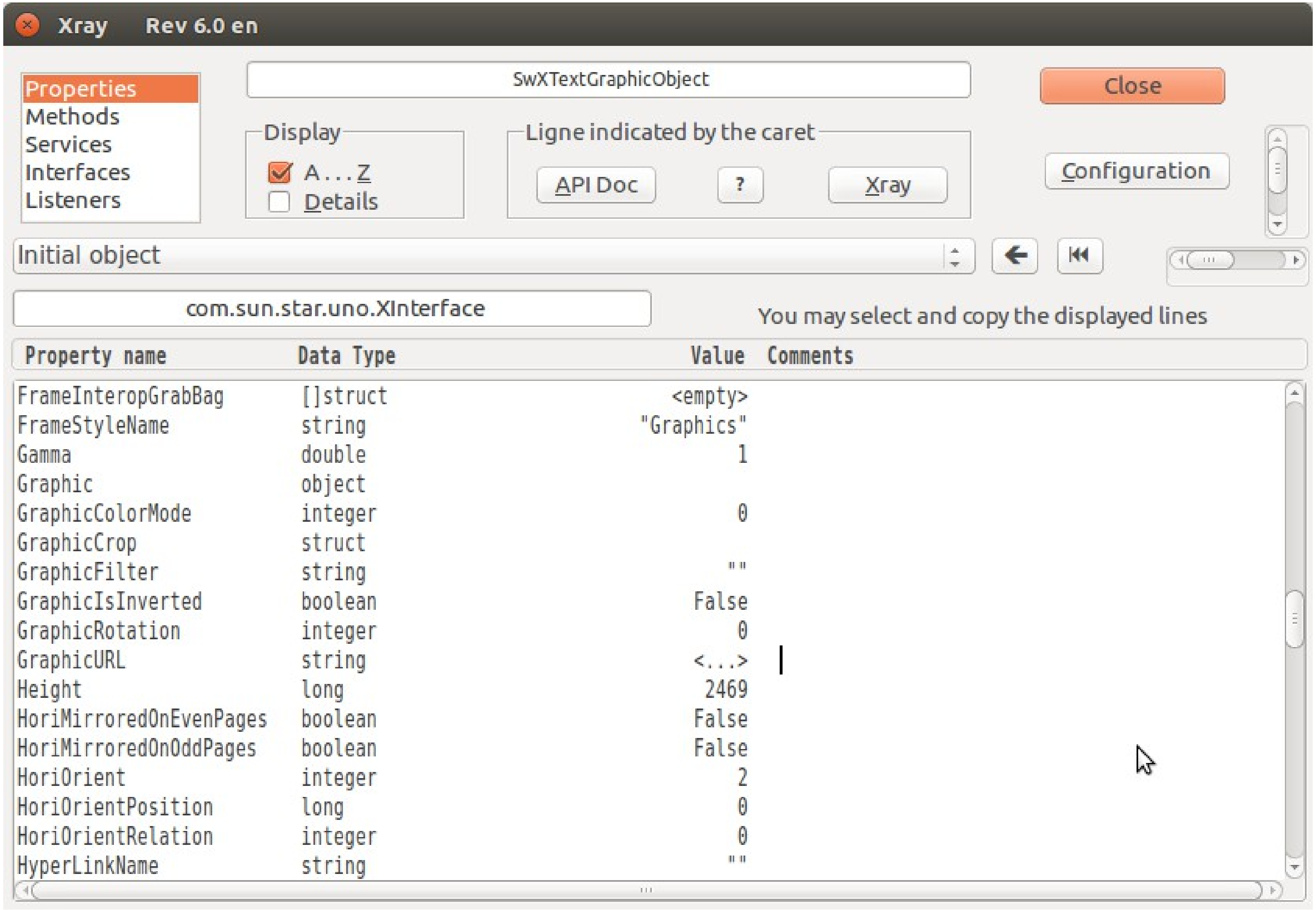
LibreOffice Hang Issues

- LibreOffice Hangs while opening/saving docx file
- Reasons can be -
  - Removed required UNO Properties
    - PROP_PARA_LINE_SPACING
    - Code reference : https://gerrit.libreoffice.org/#/c/9560
  - Not handled some required XML attributes
    - Code reference : https://gerrit.libreoffice.org/#/c/8632/
  - Memory Leaks
    - Code Reference : https://gerrit.libreoffice.org/#/c/6850

Some Useful Tools

- Xray Tool
- OOXML Tools (Chrome Browser plug-in)
- Open XML SDK Productivity tool. (for windows)

XRAY Tool



SwXTextGraphicObject

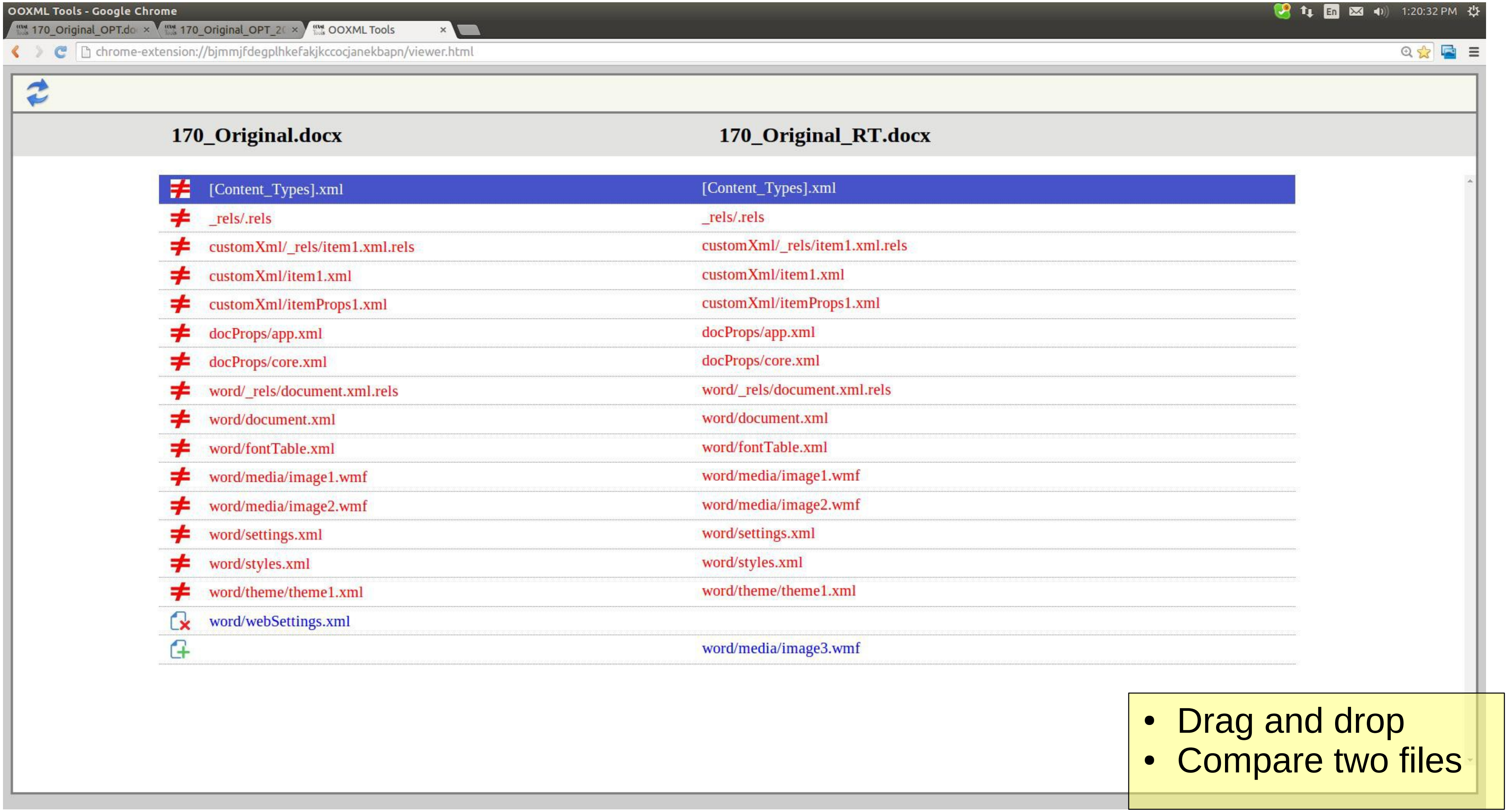
☑ A...Z
□ Details

com.sun.star.uno.XInterface

You may select and copy the displayed lines

Property name	Data Type	Value	Comments
FrameInteropGrabBag	[]struct	<empty>	
FrameStyleName	string	"Graphics"	
Gamma	double	1	
Graphic	object		
GraphicColorMode	integer	0	
GraphicCrop	struct		
GraphicFilter	string	""	
GraphicIsInverted	boolean	False	
GraphicRotation	integer	0	
GraphicURL	string	<...>	
Height	long	2469	
HoriMirroredOnEvenPages	boolean	False	
HoriMirroredOnOddPages	boolean	False	
HoriOrient	integer	2	
HoriOrientPosition	long	0	
HoriOrientRelation	integer	0	
HyperLinkName	string	""	

OOXML Tools developed by Atul Moglewar from Synerzip.



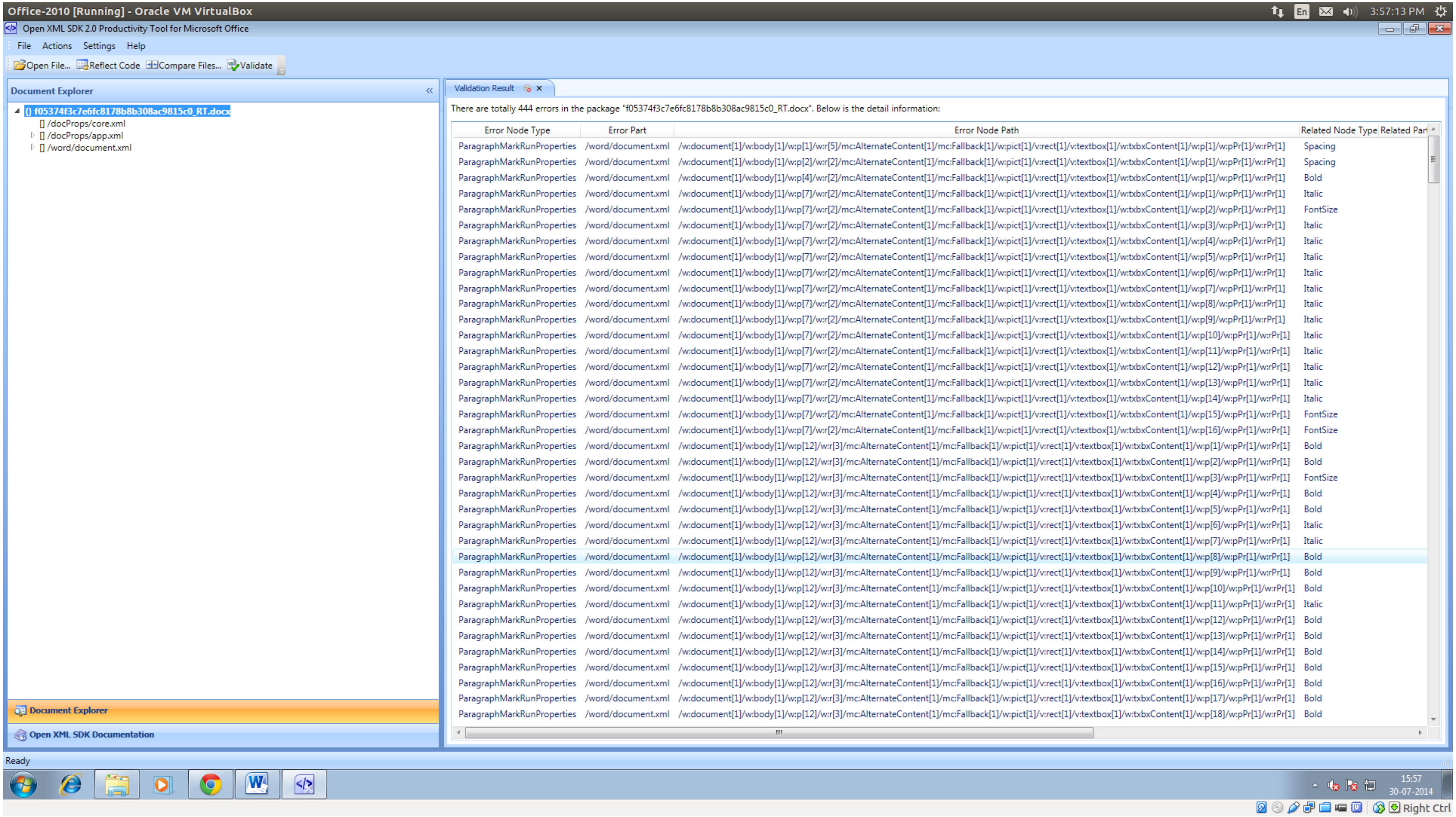
OOXML Tools - Google Chrome

chrome-extension://bjmmjfdegplhkefakjkccocjanekbapn/viewer.html

	170_Original.docx	170_Original_RT.docx
≠	[Content_Types].xml	[Content_Types].xml
≠	_rels/.rels	_rels/.rels
≠	customXml/_rels/item1.xml.rels	customXml/_rels/item1.xml.rels
≠	customXml/item1.xml	customXml/item1.xml
≠	customXml/itemProps1.xml	customXml/itemProps1.xml
≠	docProps/app.xml	docProps/app.xml
≠	docProps/core.xml	docProps/core.xml
≠	word/_rels/document.xml.rels	word/_rels/document.xml.rels
≠	word/document.xml	word/document.xml
≠	word/fontTable.xml	word/fontTable.xml
≠	word/media/image1.wmf	word/media/image1.wmf
≠	word/media/image2.wmf	word/media/image2.wmf
≠	word/settings.xml	word/settings.xml
≠	word/styles.xml	word/styles.xml
≠	word/theme/theme1.xml	word/theme/theme1.xml
✗	word/webSettings.xml	
+		word/media/image3.wmf

- Drag and drop
- Compare two files

Open SDK Tool



{} f05374f3c7e6fc8178b8b308ac9815c0_RT.docx
[] /docProps/core.xml
[] /docProps/app.xml
[] /word/document.xml

There are totally 444 errors in the package "f05374f3c7e6fc8178b8b308ac9815c0_RT.docx". Below is the detail information:

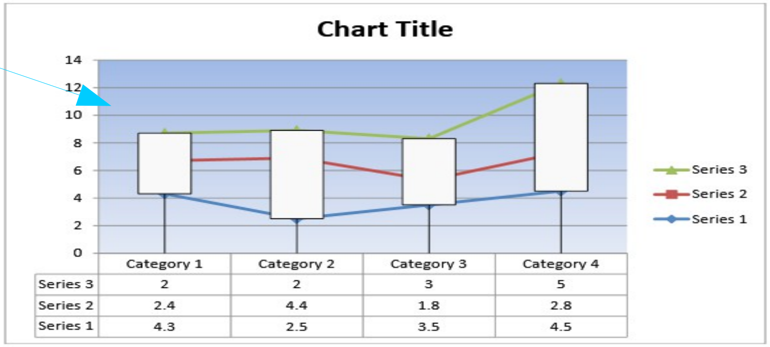
Error Node Type	Error Part	Error Node Path	Related Node Type	Related Part
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[1]/w:r[5]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[1]/w:pPr[1]/w:rPr[1]	Spacing	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[2]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[1]/w:pPr[1]/w:rPr[1]	Spacing	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[4]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[1]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[1]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[2]/w:pPr[1]/w:rPr[1]	FontSize	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[3]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[4]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[5]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[6]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[7]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[8]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[9]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[10]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[11]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[12]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[13]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[14]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[15]/w:pPr[1]/w:rPr[1]	FontSize	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[7]/w:r[2]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[16]/w:pPr[1]/w:rPr[1]	FontSize	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[1]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[2]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[3]/w:pPr[1]/w:rPr[1]	FontSize	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[4]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[5]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[6]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[7]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[8]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[9]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[10]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[11]/w:pPr[1]/w:rPr[1]	Italic	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[12]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[13]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[14]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[15]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[16]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[17]/w:pPr[1]/w:rPr[1]	Bold	
ParagraphMarkRunProperties	/word/document.xml	/w:document[1]/w:body[1]/w:p[12]/w:r[3]/mc:AlternateContent[1]/mc:Fallback[1]/w:pict[1]/v:rect[1]/v:textbox[1]/w:txbxContent[1]/w:p[18]/w:pPr[1]/w:rPr[1]	Bold	

More Examples

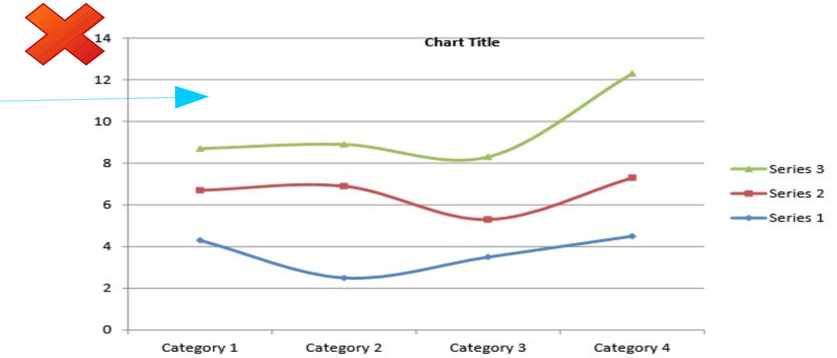
Chart

- Wall Color was missing From exported file

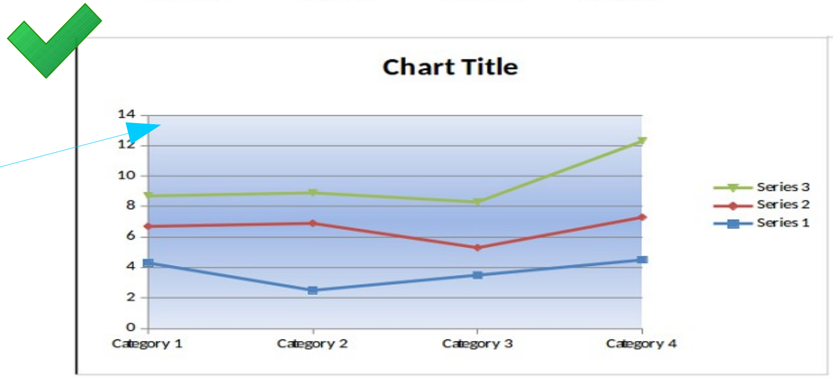
Wall color



Lost

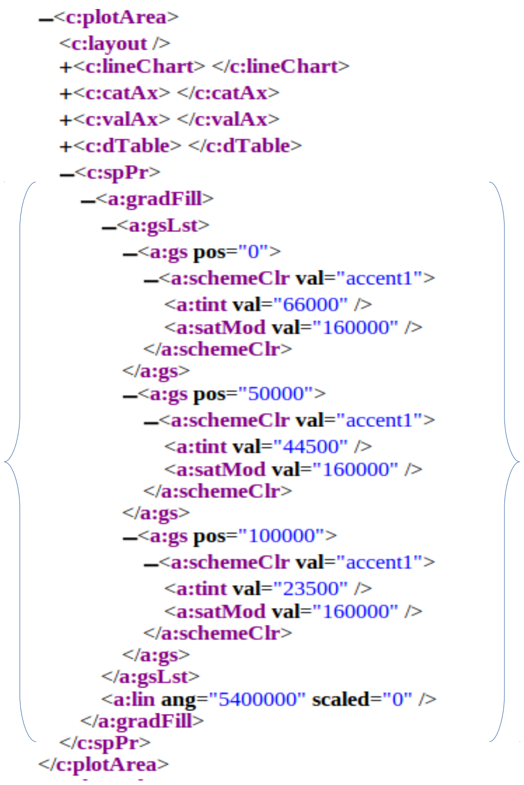


Fixed



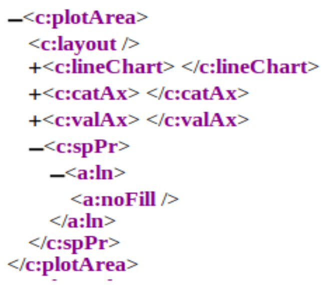
Chart

Original XML for Chart Wall Color



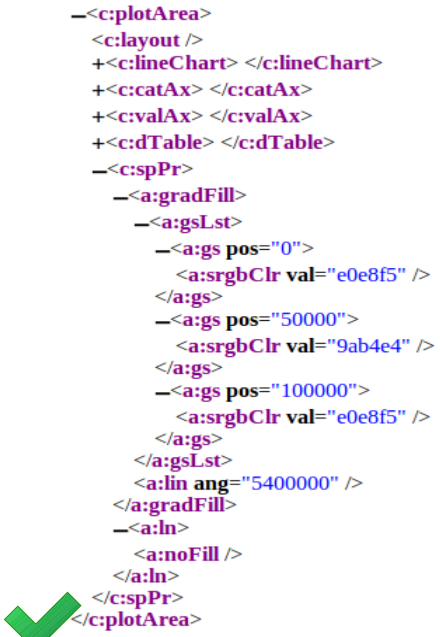
```
–<c:plotArea>
   <c:layout />
  +<c:lineChart> </c:lineChart>
  +<c:catAx> </c:catAx>
  +<c:valAx> </c:valAx>
  +<c:dTable> </c:dTable>
  –<c:spPr>
    –<a:gradFill>
      –<a:gsLst>
        –<a:gs pos="0">
          –<a:schemeClr val="accent1">
              <a:tint val="66000" />
              <a:satMod val="160000" />
           </a:schemeClr>
         </a:gs>
        –<a:gs pos="50000">
          –<a:schemeClr val="accent1">
              <a:tint val="44500" />
              <a:satMod val="160000" />
           </a:schemeClr>
         </a:gs>
        –<a:gs pos="100000">
          –<a:schemeClr val="accent1">
              <a:tint val="23500" />
              <a:satMod val="160000" />
           </a:schemeClr>
         </a:gs>
       </a:gsLst>
       <a:lin ang="5400000" scaled="0" />
     </a:gradFill>
   </c:spPr>
</c:plotArea>
```

LO : Export before fix



```
–<c:plotArea>
   <c:layout />
  +<c:lineChart> </c:lineChart>
  +<c:catAx> </c:catAx>
  +<c:valAx> </c:valAx>
  –<c:spPr>
    –<a:ln>
       <a:noFill />
     </a:ln>
   </c:spPr>
</c:plotArea>
```

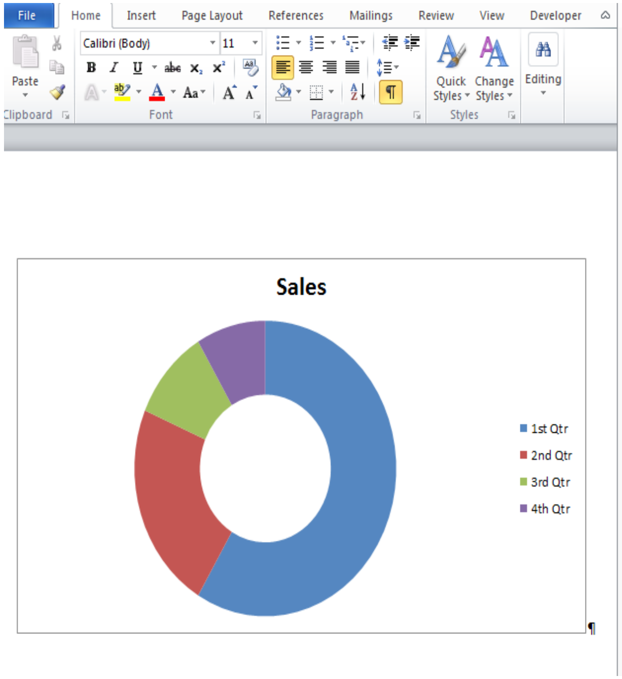
Export After Fix



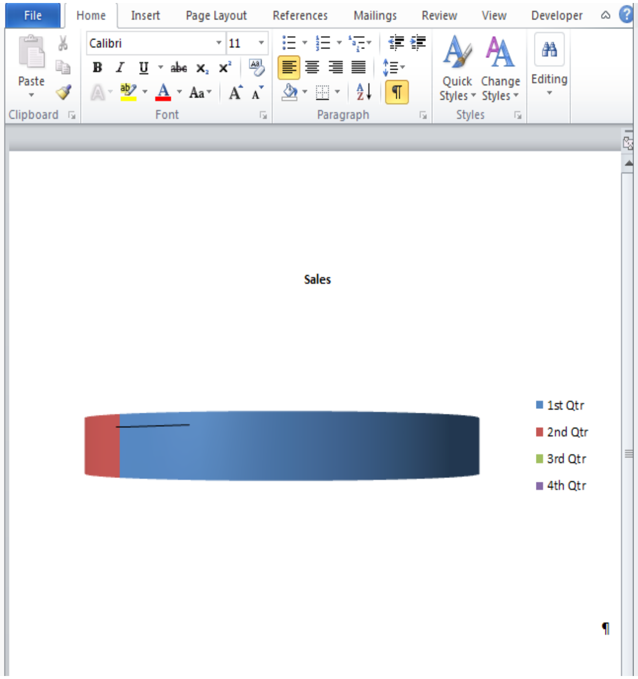
```
–<c:plotArea>
   <c:layout />
  +<c:lineChart> </c:lineChart>
  +<c:catAx> </c:catAx>
  +<c:valAx> </c:valAx>
  +<c:dTable> </c:dTable>
  –<c:spPr>
    –<a:gradFill>
      –<a:gsLst>
        –<a:gs pos="0">
            <a:srgbClr val="e0e8f5" />
         </a:gs>
        –<a:gs pos="50000">
            <a:srgbClr val="9ab4e4" />
         </a:gs>
        –<a:gs pos="100000">
            <a:srgbClr val="e0e8f5" />
         </a:gs>
       </a:gsLst>
       <a:lin ang="5400000" />
     </a:gradFill>
    –<a:ln>
       <a:noFill />
     </a:ln>
   </c:spPr>
</c:plotArea>
```

Code References : https://gerrit.libreoffice.org/7739
https://gerrit.libreoffice.org/7792

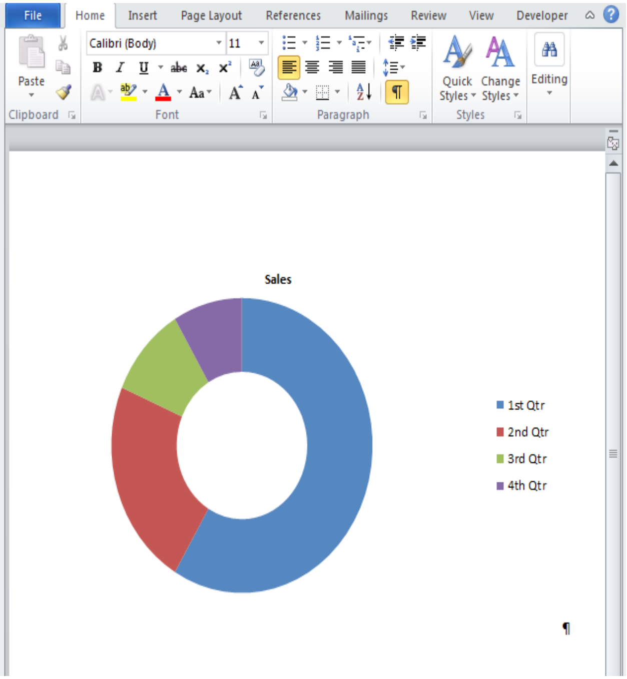
Doughnut chart



Original chart



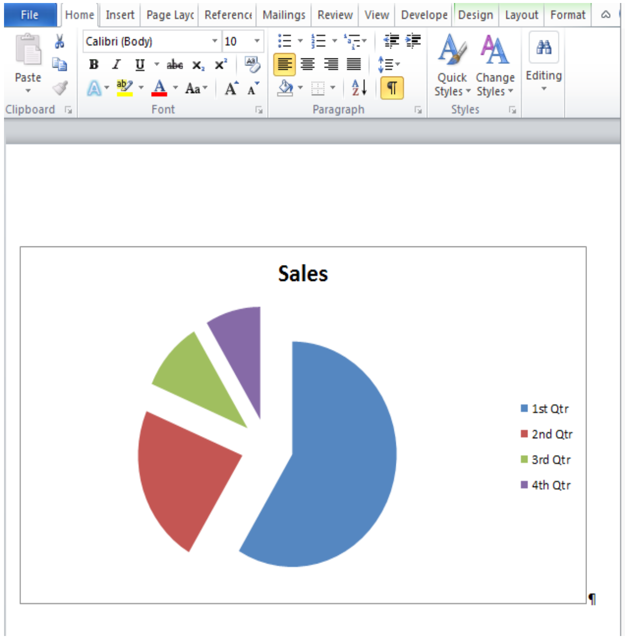
Before fix



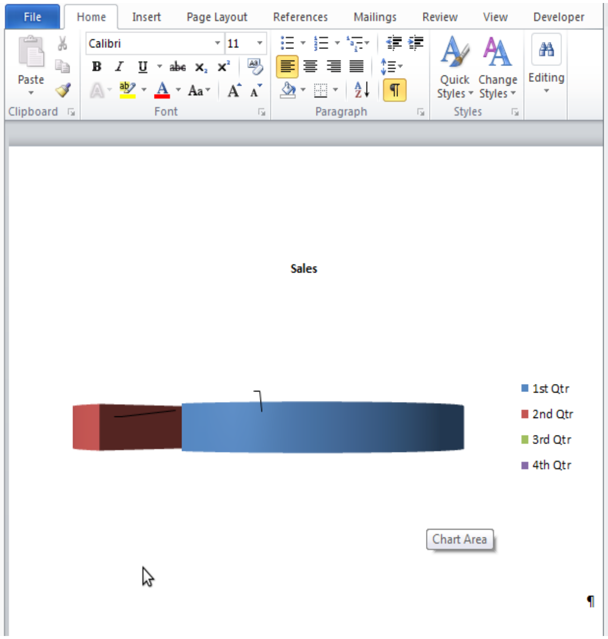
After fix

Code Reference : https://gerrit.libreoffice.org/#/c/6924

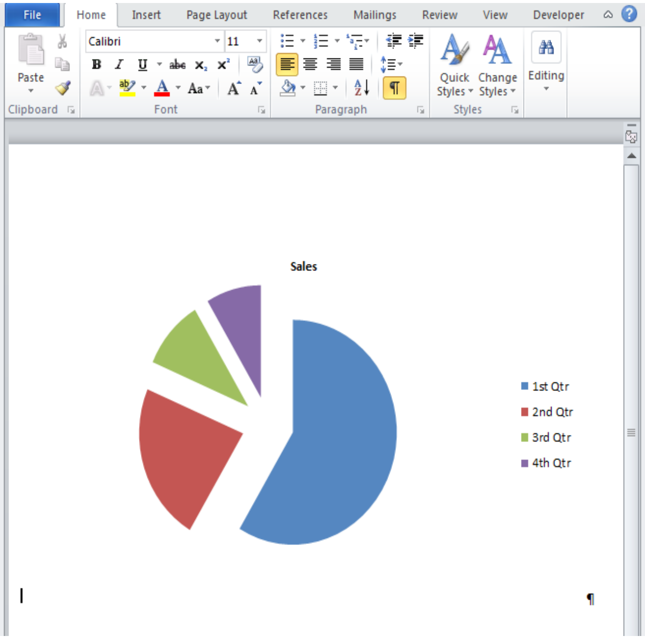
Exploded Pie Chart



Original chart



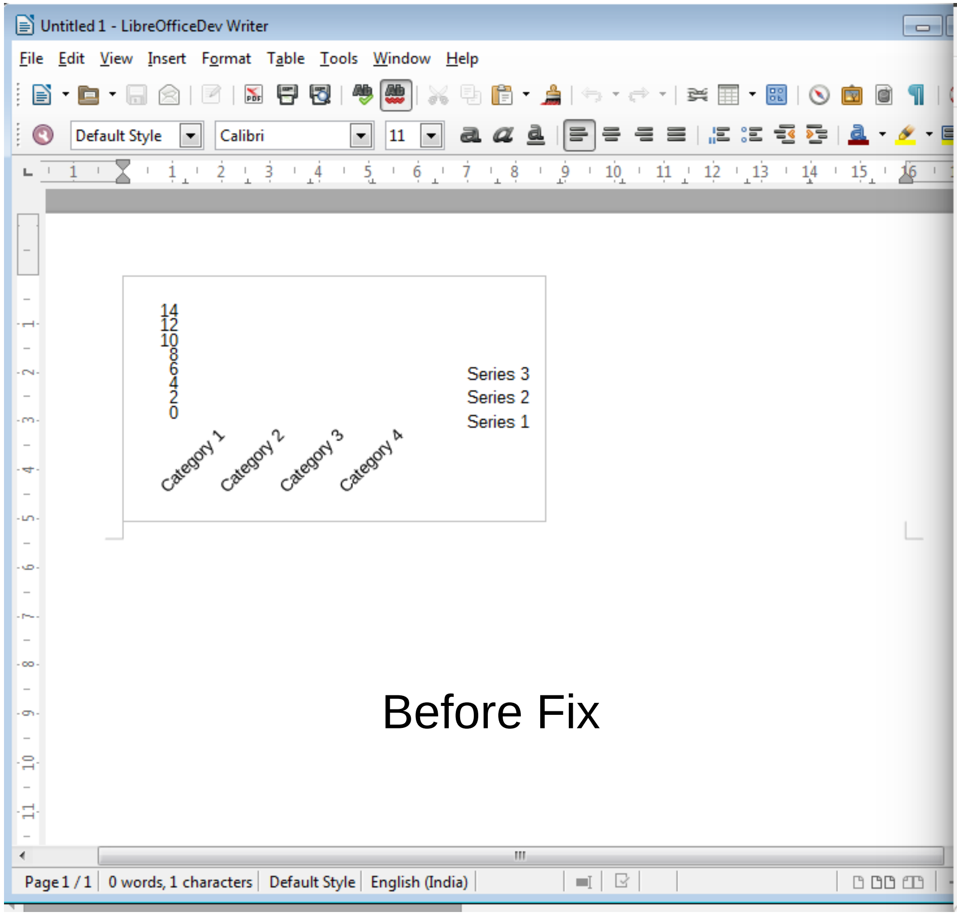
Before fix



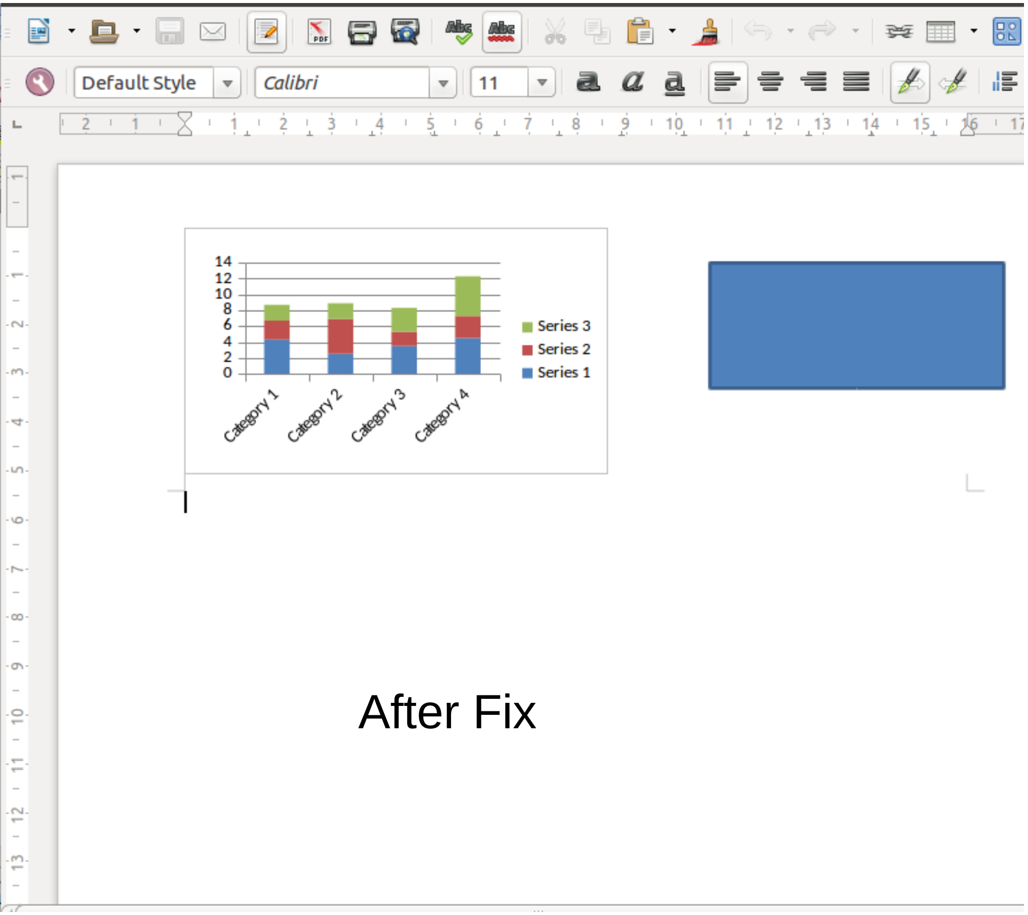
After fix

Code Reference : https://gerrit.libreoffice.org/#/c/6924

Shapes in header



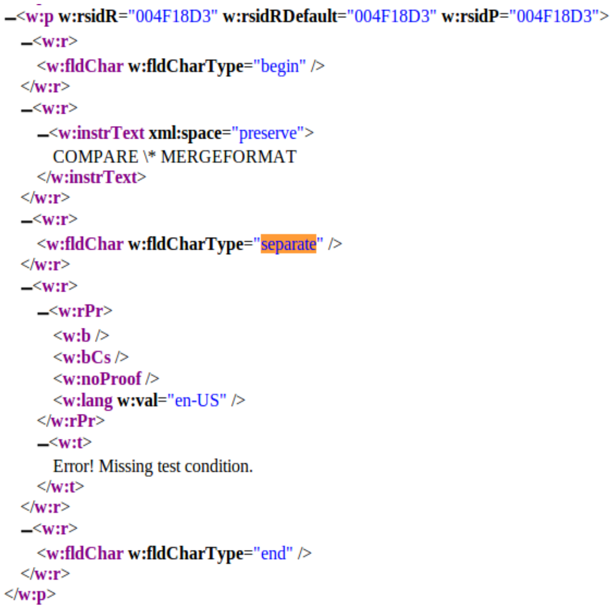
Before Fix



After Fix

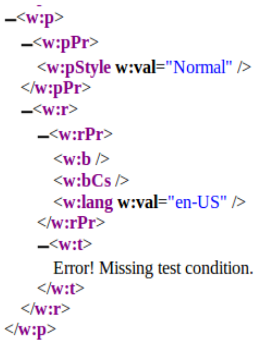
Fields

Original XML



```
–<w:p w:rsidR="004F18D3" w:rsidRDefault="004F18D3" w:rsidP="004F18D3">
  –<w:r>
     <w:fldChar w:fldCharType="begin" />
  </w:r>
  –<w:r>
     –<w:instrText xml:space="preserve">
        COMPARE \* MERGEFORMAT
     </w:instrText>
  </w:r>
  –<w:r>
     <w:fldChar w:fldCharType="separate" />
  </w:r>
  –<w:r>
     –<w:rPr>
        <w:b />
        <w:bCs />
        <w:noProof />
        <w:lang w:val="en-US" />
     </w:rPr>
     –<w:t>
        Error! Missing test condition.
     </w:t>
  </w:r>
  –<w:r>
     <w:fldChar w:fldCharType="end" />
  </w:r>
</w:p>
```

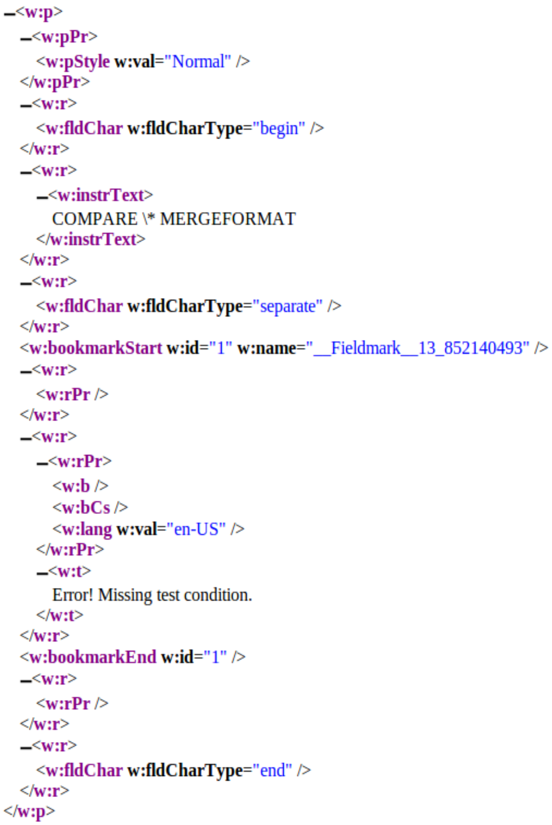
Before Fix



```
–<w:p>
  –<w:pPr>
     <w:pStyle w:val="Normal" />
  </w:pPr>
  –<w:r>
     –<w:rPr>
        <w:b />
        <w:bCs />
        <w:lang w:val="en-US" />
     </w:rPr>
     –<w:t>
        Error! Missing test condition.
     </w:t>
  </w:r>
</w:p>
```

✗

After Fix



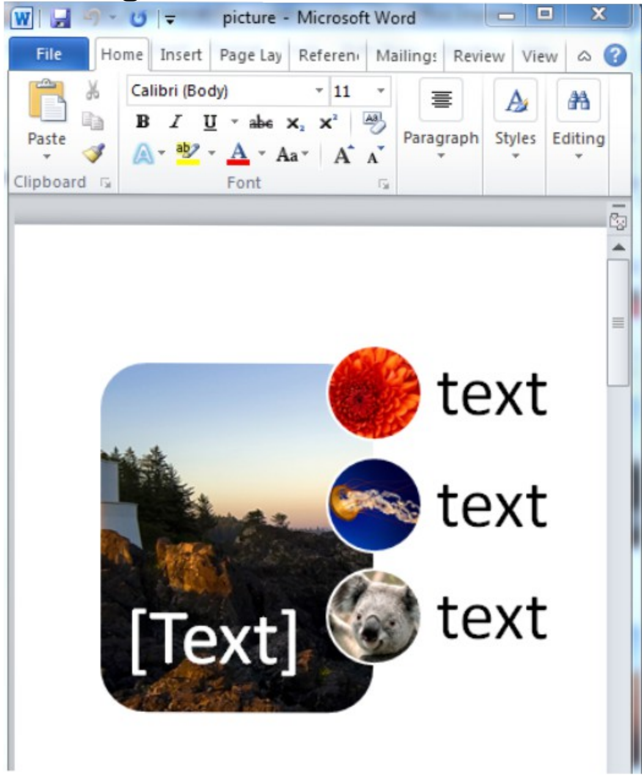
```
–<w:p>
  –<w:pPr>
     <w:pStyle w:val="Normal" />
  </w:pPr>
  –<w:r>
     <w:fldChar w:fldCharType="begin" />
  </w:r>
  –<w:r>
     –<w:instrText>
        COMPARE \* MERGEFORMAT
     </w:instrText>
  </w:r>
  –<w:r>
     <w:fldChar w:fldCharType="separate" />
  </w:r>
  <w:bookmarkStart w:id="1" w:name="__Fieldmark__13_852140493" />
  –<w:r>
     <w:rPr />
  </w:r>
  –<w:r>
     –<w:rPr>
        <w:b />
        <w:bCs />
        <w:lang w:val="en-US" />
     </w:rPr>
     –<w:t>
        Error! Missing test condition.
     </w:t>
  </w:r>
  <w:bookmarkEnd w:id="1" />
  –<w:r>
     <w:rPr />
  </w:r>
  –<w:r>
     <w:fldChar w:fldCharType="end" />
  </w:r>
</w:p>
```

✓

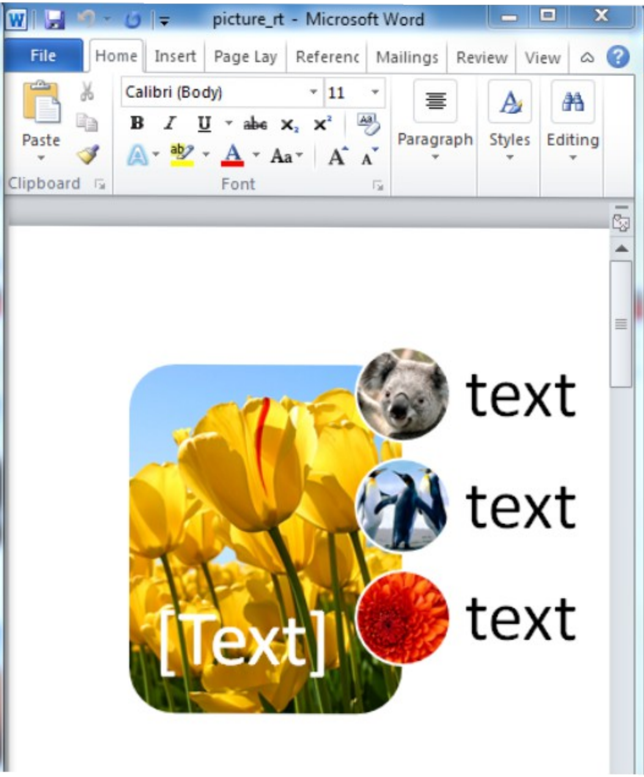
Smart Art

Image Fills in smart are exported properly.

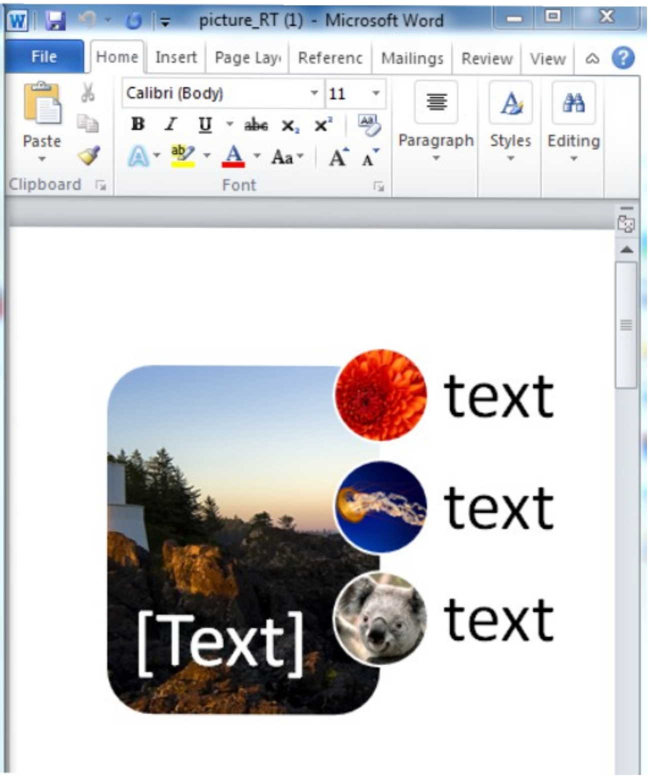
Original File



LO Export : Before Fix



After Fix



Code reference : https://gerrit.libreoffice.org/#/c/9121

Synerzip's Contribution

- ~250 patches submitted by synerzip in last 1 year.
- 50+ scenarios of crash/corruption fixed.
- 270+ bugs filed on BugZilla.
- 200+ bugs resolved.

Team Synerzip

References

- http://cgit.freedesktop.org/libreoffice/core/log/?qt=author&q=synerzip
- http://msdn.microsoft.com/en-us/library/office/gg607163(v=office.14).aspx
- http://www.ecma-international.org/publications/standards/Ecma-376.htm
- http://www.datypic.com/sc/ooxml/
- https://chrome.google.com/webstore/detail/ooxml-tools/bjmmjfdegplhkefakjkccocjanekbapn?hl=en-US&utm_source=chrome-ntp-launcher
- https://wiki.documentfoundation.org/Macros

Thank You.