Office Binary (doc, xls, ppt) Translator to Open XML

(xls to xlsx Mapping v0.3; DlaLOGIKa/sare/wk 27 March 2009)

This is a very initial mapping table from the binary Excel format (xls) to SpreadsheetML -- the Open XML markup language for Excel documents (xlsx). We will use this mapping table in the future to implement the converter from binary to Open XML Excel documents.

The following mapping information is currently available:

- Document Parts
- Workbook Part
- SharedStrings Part
- Worksheet Part
- External Link Part
- Styles

The following columns describe the mapping:

Stream: The binary stream **Struct:** The binary structure

Field: The mapped field of the binary structure

XML File: The corresponding XML fileXML Element: The corresponding XML elementXML Attribute: The corresponding XML attribute

Comment: *Interesting details*

| Document Parts | | | |
|-----------------------|---|--|-------------|
| Stream | XML file | Comment | Implemented |
| Workbook | xl\workbook.xml | Generel information about the workbook and the worksheets. | ✓ |
| | | There's a list with worksheets. | |
| Workbook | xl\sharedStrings.xml | Global string table. Includes every string within the excel | ✓ |
| | | document. | |
| Workbook | xl\styles.xml | This file includes every style format. | ✓ |
| Workbook | xl\worksheets\sheetXYZ.xml | There are several worksheetfiles, every worksheet or diagram | ✓ |
| | | has its own <name>.xml . These sheets are listed in the</name> | |
| | | workbook.xml. | |
| Workbook | xl\externalLinks\externalLinkXYZ.xml | There are several externalLink files. For every external | ✓ |
| | | reference there exists an external link file. | |
| Workbook | xl\externalLinks_rels\externalLinkXYZ.xml.rels | This file includes the path to the external file. | ✓ |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

| Workbook Part BiffRecord | Field | XML element | XML attribute | Comment | Implemented |
|--------------------------|------------|--|-----------------|---|-------------|
| BOUNDSHEET | rgch | /workbook/sheets/sheet | name | This is a byte array which contains the name of a worksheet | ✓ |
| | | /workbook/sheets/sheet | sheetId | Automatic generated sheet id | ✓ |
| | | /workbook/sheets/sheet | r:Id | Automatic generated sheet id and an additional string | ✓ |
| SUPBOOK | | /externalReferences/externalReference/ | | For each external link file there is a entry in the externalReferences element | √ |
| SUPBOOK | | /externalReferences/externalReference/ | r:ld | Automatic generated external link reference id | ✓ |
| NAME / Lbl | | /workbook/definedNames/definedName/ | | Defined name structure | ✓ |
| | rgce + cce | /workbook/definedNames/definedName/ | | Formula for this defined name, the cce value defines the | ✓ |
| | Name + cch | /workbook/definedNames/definedName/ | name | lenght of the parsed formula in the rgce field The cch field defines the lenght of the defined name string | ✓ |
| | itab | /workbook/definedNames/definedName/ | localSheetId | This is the reference to a worksheet | ✓ |
| | chKey | /workbook/definedNames/definedName/ | shortcutKey | Shortcut to this name | |
| | fGrp | /workbook/definedNames/definedName/ | functionGroupId | ID from a functionGroup | |
| NAMECMT | comment | /workbook/definedNames/definedName/ | comment | Comment to a defined name | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

| | etala | VAAL alamant | V8.61 -4414 | Community | |
|----------|-------------------------------------|--------------------------|--------------------|--|-------------|
| ffRecord | Field | XML element | XML attribute | Comment | Implemented |
| ST | cstTotal | /sst | count | This attribute is used to show how many references are on | • |
| | | | | these shared strings. | |
| ST | cstUnique | /sst | uniqueCount | How many unique shared strings are available. | √ |
| Т | rest of the SST struct (BiffRecord) | /sst/si/r/t | | Text from a unique string. These strings have their own format | ✓ |
| | | | | within the rest of the SharedStringTable BiffRecord. Every | |
| | | | | String contains of a length value and an option bitmask. The | |
| | | | | bitmask is used differ how to read the rest of the byte array. | |
| IT | | \sst\si\r\rPr\ | | Font element | ✓ |
| | dyHeight | \sst\si\r\rPr\sz\ | val | Font size | ✓ |
| | fItalic | \sst\si\r\rPr\i\ | | Is italic | ✓ |
| | fStrikeOut | \sst\si\r\rPr\strike\ | val | | ✓ |
| | fOutline | \sst\si\r\rPr\outline\ | val | | ✓ |
| | fShadow | \sst\si\r\rPr\shadow\ | val | | ✓ |
| | fCondense | \sst\si\r\rPr\condense\ | val | | ✓ |
| | icv | \sst\si\r\rPr\color\ | auto, rgb, indexed | The color is always the rgb value! | ✓ |
| | bls | \sst\si\r\rPr\b\ | val | bolt | ✓ |
| | SSS | \sst\si\r\rPr\vertAlign\ | val | | ✓ |
| | uls | \sst\si\r\rPr\u\ | val | underline | ✓ |
| | bFamily | \sst\si\r\rPr\family\ | val | | ✓ |
| | bCharSet | \sst\si\r\rPr\charset\ | val | | ✓ |
| | fontName | \sst\si\r\rPr\rFont\ | val | | ✓ |

| Worksheet Part | | | | | |
|--|--|---------------------------------|---------------|---|-------------|
| iffRecord | Field | XML element | XML attribute | Comment | Implemented |
| IUMBER, RK, IULRK, BLANK, IULBLANK, ABELSST | rw | /worksheet/sheetData/row | r | Indicates where the following data row is located | ✓ |
| NUMBER, RK, MULRK, BLANK, MULBLANK, ABELSST | Different fields: NUMBER, RK, BLANK and LABELSST use the col value. The MULBLANK and MULRK structs use the colFirst value + an integer value that indicates the position of the array. | /worksheet/sheetData/row/c | r | Indicates where the following data is located. This value has to be a mix of the column as letter and the row as a number value. | ✓ |
| | | /worksheet/sheetData/row/c | t | This value indicates if the following data is a string or a number. If it is a number, the attribute "t" is missing else the attribute t is there and its value is "s". | ✓ |
| | ixfe | /worksheet/sheetData/row/c | S | Reference to the style structure | ✓ |
| LABELSST | isst | /worksheet/sheetData/row/c/v | | The data. If the value is a string then it is only the reference to the shared string table. | ✓ |
| MULRK | rgrkrec -> array of rk records | /worksheet/sheetData/row/c/v | | 11 | ✓ |
| MULBLANK, BLANK | no value | /worksheet/sheetData/row/c/v | | н | ✓ |
| RK | rk | /worksheet/sheetData/row/c/v | | 11 | ✓ |
| NUMBER | num | /worksheet/sheetData/row/c/v | | 11 | ✓ |
| MERGECELL | cmcs | /worksheet/mergeCells | count | | ✓ |
| MERGECELL | rwFirst, rwLast, colFirst, colLast | /worksheet/mergeCells/mergeCell | ref | Creates a reference string that has two values, the position of the top left corner from the mergecell and the bottom right position | ✓ |
| FORMULA | rw, col | /worksheet/sheetData/row/c | r | Indicates where the following data is located. This value has to be a mix of the column as letter and the row as a number value. | ✓ |
| | cce, rgce | /worksheet/sheetData/row/c/f | | This field is used to store the whole formula. This formula is stored in the postfix notation and has to be translated into the infix notation. | ✓ |
| | fShrFmla | /worksheet/sheetData/row/c/f | t="shared" | If this formula is a shared formula this flag is set | ✓ |
| | val | /worksheet/sheetData/row/c/v | | The calculated value from a formula, this can be a number, a string, an error value or a boolean. | ✓ |
| | | /worksheet/sheetData/row/c/f | si | Shared group - automatically generated number | ✓ |
| ARRAY | rwFirst, rwLast, colFirst,colLast | /worksheet/sheetData/row/c | r | Indicates where the following data is located. This value has to be a mix of the column as letter and the row as a number value. | √ |
| | cce, rgce | /worksheet/sheetData/row/c/f | | This field is used to store the whole formula. This formula is stored in the postfix notation and has to be translated into the infix notation. | ✓ |
| | | /worksheet/sheetData/row/c/f | t="array" | Indicates that the formula is an array formula | ✓ |
| | | /worksheet/sheetData/row/c/f | ref=" " | The reference value is the same value where the formula is located | ✓ |
| HRFMLA | ref | /worksheet/sheetData/row/c/f | ref | Reference where this shared formula is positioned | ✓ |
| TRING | | /worksheet/sheetData/row/c/v | | This record specifies the string value of a formula. | ✓ |
| ROW | fCollapsed | /worksheet/sheetData/row | collapsed | If the row is collapsed this is true. | ✓ |
| | fGhostDirty | /worksheet/sheetData/row | customFormat | | ✓ |
| | fUnsynced | /worksheet/sheetData/row | customHeight | | ✓ |

| Worksheet Part | | | | | |
|----------------|-----------------------|---|------------------|---|-------------|
| iffRecord | Field | XML element | XML attribute | Comment | Implemented |
| | fDyZero | /worksheet/sheetData/row | hidden | Hiddenflag | ✓ |
| | miyRw | /worksheet/sheetData/row | ht | Hight of this row | ✓ |
| | iOutLevel | /worksheet/sheetData/row | outlineLevel | Specifies the outline level | ✓ |
| | fPhonetic | /worksheet/sheetData/row | ph | phonetic data flag | ✓ |
| | ixfe_val | /worksheet/sheetData/row | S | A reference to the style data element | ✓ |
| | colMic, colMac | /worksheet/sheetData/row | spans | The area in which are values | ✓ |
| | fExDes | /worksheet/sheetData/row | thickBot | If the row above has a thick bottom border this value is true | ✓ |
| | fExAsc | /worksheet/sheetData/row | thickTop | If the row below has a thick top border this value is true | ✓ |
| OLINFO | | /worksheet/cols/col Defines some column information | | Defines some column information | ✓ |
| | colFirst | /worksheet/cols/col | min | First cell in this column | ✓ |
| | colLast | /worksheet/cols/col | max | Last cell in this column | ✓ |
| | coldx | /worksheet/cols/col | width | The widt from this column | ✓ |
| | fUserSet | /worksheet/cols/col | customWidth | User set the widht! | ✓ |
| | fHidden | /worksheet/cols/col | hidden | If the column is hidden this value is 1 | ✓ |
| | fBestFit | /worksheet/cols/col | bestFit | bestFit flag | ✓ |
| | fPhonetic | /worksheet/cols/col | phonetic | phonetic data flag | ✓ |
| | iOutLevel | /worksheet/cols/col | outlineLevel | outlineLevel value | ✓ |
| | fCollapsed | /worksheet/cols/col | collapsed | collapsed flag | ✓ |
| | ixfe | /worksheet/cols/col | style | ID to the style record | ✓ |
| | | | · | | |
| EFCOLWIDTH | cchdefColWidth | /worksheet/sheetFormatPr | defaultColWidth | Default column width | ✓ |
| DEFROWHEIGHT | miyRW | /worksheet/sheetFormatPr | defaultRowHeight | Default row hight | ✓ |
| | fDyZero | /worksheet/sheetFormatPr | zeroHeight | Flag is set if the row height is 0 | ✓ |
| | fUnsynced | /worksheet/sheetFormatPr | customHeight | If the height is set manually | ✓ |
| | fExAsc | /worksheet/sheetFormatPr | thickTop | If the row above has a thick bottom border this value is true | ✓ |
| | fExDsc | /worksheet/sheetFormatPr | thickBottom | If the row below has a thick top border this value is true | ✓ |
| | | | | | |
| TUP | | /worksheet/pageSetup/ | | This structure stores the page setup information | ✓ |
| | iPaperSize | /worksheet/pageSetup/ | paperSize | | ✓ |
| | iScale | /worksheet/pageSetup/ | scale | | ✓ |
| | fNoColor | /worksheet/pageSetup/ | blackAndWhite | | ✓ |
| | iCopies | /worksheet/pageSetup/ | copies | | ✓ |
| | fDraft | /worksheet/pageSetup/ | draft | | ✓ |
| | iPageStart | /worksheet/pageSetup/ | firstPageNumber | | ✓ |
| | iFitHeight | /worksheet/pageSetup/ | fitToHeight | | ✓ |
| | iFitWidth | /worksheet/pageSetup/ | fitToWidth | | ✓ |
| | iRes | /worksheet/pageSetup/ | horizontalDpi | | ✓ |
| | fUsePage | /worksheet/pageSetup/ | useFirstPageNumb | | ✓ |
| | iVRes | /worksheet/pageSetup/ | verticalDpi | | ✓ |
| | iError | /worksheet/pageSetup/ | errors | | ✓ |
| | fNoOrient, fLandscape | /worksheet/pageSetup/ | orientation | Page orientation, none, landscape or portrait | ✓ |
| | fNotes, fEndNotes | /worksheet/pageSetup/ | cellComments | | ✓ |
| | fLefttoRight | /worksheet/pageSetup/ | pageOrder | | ✓ |
| | <u> </u> | | , 3 | | |
| TUP | Setup.numHdr | /worksheet/pageMargins/ | header | Size of the header | ✓ |
| | Setup.numFtr | /worksheet/pageMargins/ | footer | Size of the footer | ✓ |
| FTMARGIN | | /worksheet/pageMargins/ | left | | ✓ |
| IGHTMARGIN | | /worksheet/pageMargins/ | right | | ✓ |
| C. / | | / Worksheed pagervial Bills/ | i ibiit | | |

| Worksheet Part | Worksheet Part | | | | | |
|-----------------------|----------------|-------------------------|---------------|---------|-------------|--|
| BiffRecord | Field | XML element | XML attribute | Comment | Implemented | |
| TOP | | /worksheet/pageMargins/ | top | | ✓ | |
| BOTTOM | | /worksheet/pageMargins/ | bottom | | ✓ | |
| | | | | | | |
| | | | | | | |
| | | | | | | |
| | | | | | | |

| BiffRecord | Field | XML element | XML attribute | Comment | Implemented |
|------------|-------------------|---|---------------|---|-------------|
| SUPBOOK | rgst | \externalLink\externalBook\sheetNames\sheetName | var | This field stores every worksheet name from the external | ✓ |
| | | | | workbook. For each name in the rgst field there's a new | |
| | | | | sheetName element with the var attribute. | |
| EXTERNNAME | body | \externalLink\externalBook\definedNames\definedName\ | name | This field stores self defined names from external worksheets. | ✓ |
| СТ | | \externalLink\externalBook\sheetDataSet\sheetData\ | | | ✓ |
| | | \externalLink\externalBook\sheetDataSet\sheetData\ | sheetId | Generated id | ✓ |
| CRN | | \externalLink\externalBook\sheetDataSet\sheetData\row\ | | | ✓ |
| CRN | rw | \externalLink\externalBook\sheetDataSet\sheetData\row\ | r | The row value from this external reference | ✓ |
| CRN | colFirst, colLast | \externalLink\externalBook\sheetDataSet\sheetData\row\cell\ | | Create cell elements for each value between colFirst and | ✓ |
| | | | | colLast | |
| CRN | oper | \externalLink\externalBook\sheetDataSet\sheetData\row\cell\v\ | | The oper value stores the value from each cell | ✓ |
| CRN | oper | \externalLink\externalBook\sheetDataSet\sheetData\row\cell\ | t | This attribute shows the type of the cell, this can be a boolean, | ✓ |
| | | | | a string or a number value. | |

| External Link Rels Part | | | | | |
|-------------------------|----------|------------------------------|---------------|---|-------------|
| BiffRecord | Field | XML element | XML attribute | Comment | Implemented |
| SUPBOOK | virtPath | /Relationships/Relationship/ | Target | The path and filename to the external excel file. | ✓ |
| SUPBOOK | | /Relationships/Relationship/ | TargetMode | This attribute must be "External" | ✓ |

| Styles Part | | | | | |
|-------------|-------------|---|--------------------|--|-------------|
| BiffRecord | Field | XML element | XML attribute | Comment | Implemented |
| | | \styleSheet\numFmts\ | count | Number of format elements | ✓ |
| FORMAT | ifmt | \styleSheet\numFmts\numFmt\ | numFmtld | The ID from this format string | ✓ |
| | stFormat | \styleSheet\numFmts\numFmt\ | formatCode | The string from this format | ✓ |
| | | \styleSheet\fonts\ | count | Number of font elements | ✓ |
| FONT | | \styleSheet\fonts\font\ | | Font element | ✓ |
| | dyHeight | \styleSheet\fonts\font\sz\ | val | Font size | ✓ |
| | fItalic | \styleSheet\fonts\font\i\ | | Is italic | ✓ |
| | fStrikeOut | \styleSheet\fonts\font\strike\ | val | | ✓ |
| | fOutline | \styleSheet\fonts\font\outline\ | val | | ✓ |
| | fShadow | \styleSheet\fonts\font\shadow\ | val | | ✓ |
| | fCondense | \styleSheet\fonts\font\condense\ | val | | ✓ |
| | icv | \styleSheet\fonts\font\color\ | auto, rgb, indexed | The color is always the rgb value! | ✓ |
| | bls | \styleSheet\fonts\font\b\ | val | | ✓ |
| | SSS | \styleSheet\fonts\font\vertAlign\ | val | | ✓ |
| | uls | \styleSheet\fonts\font\u\ | val | | ✓ |
| | bFamily | \styleSheet\fonts\font\family\ | val | | ✓ |
| | bCharSet | \styleSheet\fonts\font\charset\ | val | | ✓ |
| | fontName | \styleSheet\fonts\font\name\ | val | | ✓ |
| | | | | | |
| STYLE | rgch | \styleSheet\cellStyles\cellStyle\ | name | Name of the style. | ✓ |
| | ixfe | \styleSheet\cellStyles\cellStyle\ | xfld | ID to a xf element | ✓ |
| | istyBuiltIn | \styleSheet\cellStyles\cellStyle\ | builtinId | | ✓ |
| | | | | | |
| | | \styleSheet\fills\ | count | | ✓ |
| XF | fls | \styleSheet\fills\fill\patternFill\ | patternType | | ✓ |
| | icvFore | \styleSheet\fills\fill\patternFill\fgColor\ | rgb | rgb color value if the icvFore is not auto | ✓ |
| | | \styleSheet\fills\fill\patternFill\fgColor\ | auto="1" | Auto color if the icvFore is auto | ✓ |
| | icvBack | \styleSheet\fills\fill\patternFill\bgColor\ | rgb | rgb color value if the icvBack is not auto | ✓ |
| | | \styleSheet\fill\patternFill\bgColor\ | auto="1" | Auto color if the icvBack is auto | ✓ |
| | | \styleSheet\borders\ | count | Number of borderelements | ✓ |
| XF | grbitDiag | \styleSheet\borders\border\ | diagonalDown="1" | If grbit == 1 | ✓ |
| | grbitDiag | \styleSheet\borders\border\ | diagonalUp="1" | If grbit == 2, if grbit == 3 -> booth values | ✓ |
| | 5 4 4 5 | \styleSheet\borders\border\left\ | | , , , , , , , , , , , , , , , , , , , | ✓ |
| | dgLeft | \styleSheet\borders\border\left\ | style | Style from this borderpart | ✓ |
| | icvLeft | \styleSheet\borders\border\left\color\ | rgb or auto | | ✓ |
| | 1012010 | \styleSheet\borders\border\right\ | . 8 | | ✓ |
| | dgRight | \styleSheet\borders\border\right\ | style | Style from this borderpart | ✓ |
| | icvRight | \styleSheet\borders\border\right\color\ | rgb or auto | 0.5,1.0 0 0 0 0 0 0 0 | ✓ |
| | - 0 - | \styleSheet\borders\border\top\ | 5 | | ✓ |
| | dgTop | \styleSheet\borders\border\top\ | style | Style from this borderpart | ✓ |
| | icvTop | \styleSheet\borders\border\top\color\ | rgb or auto | , | ✓ |
| | | \styleSheet\borders\border\bottom\ | . 50 0. 00.0 | | ✓ |
| | dgBottom | \styleSheet\borders\border\bottom\ | style | Style from this borderpart | ✓ |
| | icvBottom | \styleSheet\borders\border\bottom\color\ | rgb or auto | 2.7.0 o o bo. do. pui c | ✓ |
| | | | | | |

| fRecord | Field | XML element | XML attribute | Comment | Implemented |
|---------|--------------|--|-----------------------|----------------------------|---------------|
| | dgDiag | \styleSheet\borders\border\diagonal\ | style | Style from this borderpart | implemented ✓ |
| | icvDiag | \styleSheet\borders\border\diagonal\color\ | rgb or auto | Style from this borderpart | · · |
| | icvDiag | /stylesneet/borders/border/diagonar/color/ | rgb or auto | | • |
| | | \styleSheet\cellStyleXfs\ | count | number of cellStyleXfs | ✓ |
| | ifmt | \styleSheet\cellStyleXfs\xf\ | numFmtId | Format ID | ✓ |
| | ifnt | \styleSheet\cellStyleXfs\xf\ | fontId | Font ID | ✓ |
| | ·····C | \styleSheet\cellStyleXfs\xf\ | filld | Autogenerated | ✓ |
| | | \styleSheet\cellStyleXfs\xf\ | borderId | Autogenerated | ✓ |
| | fWrap | \styleSheet\cellStyleXfs\xf\alignment\ | wrapText="1" | , latobene latea | ✓ |
| | alc | \styleSheet\cellStyleXfs\xf\alignment\ | horizontal | | ✓ |
| | alcV | \styleSheet\cellStyleXfs\xf\alignment\ | vertical | | ✓ |
| | fJustLast | \styleSheet\cellStyleXfs\xf\alignment\ | justifyLastLine="1" | | ✓ |
| | fShrinkToFit | \styleSheet\cellStyleXfs\xf\alignment\ | shrinkToFit="1" | | ✓ |
| | trot | \styleSheet\cellStyleXfs\xf\alignment\ | textRotation | | ✓ |
| | cIndent | \styleSheet\cellStyleXfs\xf\alignment\ | indent | | ✓ |
| | iReadOrder | \styleSheet\cellStyleXfs\xf\alignment\ | readingOrder | | ✓ |
| | | (and the second form of the fem British | | | |
| | | \styleSheet\cellXfs\ | count | Number of cellXfs | ✓ |
| | ifmt | \styleSheet\cellXfs\xf\ | numFmtld | Format ID | ✓ |
| | ifnt | \styleSheet\cellXfs\xf\ | fontId | Font ID | ✓ |
| | | \styleSheet\cellXfs\xf\ | fillId | Autogenerated | ✓ |
| | | \styleSheet\cellXfs\xf\ | borderId | Autogenerated | ✓ |
| | ixfParent | \styleSheet\cellXfs\xf\ | xfld | | ✓ |
| | ifmt | \styleSheet\cellXfs\xf\ | applyNumberFormat="1" | if Format != 0 | ✓ |
| | | \styleSheet\cellXfs\xf\ | applyBorder="1" | if Border != 0 | ✓ |
| | | \styleSheet\cellXfs\xf\ | applyFill="1" | if Fill != 0 | ✓ |
| | ifnt | \styleSheet\cellXfs\xf\ | applyFont="1" | if Font != 0 | ✓ |
| | fWrap | \styleSheet\cellXfs\xf\alignment\ | wrapText="1" | | ✓ |
| | alc | \styleSheet\cellXfs\xf\alignment\ | horizontal | | ✓ |
| | alcV | \styleSheet\cellXfs\xf\alignment\ | vertical | | ✓ |
| | fJustLast | \styleSheet\cellXfs\xf\alignment\ | justifyLastLine="1" | | ✓ |
| | fShrinkToFit | \styleSheet\cellXfs\xf\alignment\ | shrinkToFit="1" | | ✓ |
| | trot | \styleSheet\cellXfs\xf\alignment\ | textRotation | | ✓ |
| | cIndent | \styleSheet\cellXfs\xf\alignment\ | indent | | ✓ |
| | iReadOrder | \styleSheet\cellXfs\xf\alignment\ | readingOrder | | ✓ |

| Styles Part BiffRecord | | | | |
|---------------------------|-------|-------------|-----------------------|-------------|
| BiffRecord | Field | XML element | XML attribute Comment | Implemented |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |