



Collabora Productivity

OOXML support maintenance

By Tamás Zolnai

Software Engineer at Collabora Productivity
tamas.zolnai@collabora.com

Introduction

Why OOXML compatibility is an issue?

- Migration or mixed environment
- Replacement of MSO

What will I speak about?

- Areas of comp. issues
- Improvements done in last year
- Future possibilities

Form controls

Form control example

The screenshot displays the LibreOffice Writer interface with a document titled "UP5_Umfrage.docx". The document content is as follows:

1.1 Heutige Nutzung des UP5

Wie häufig nutzen Sie den UP5?

- wöchentlich
- monatlich
- 1-2mal pro Jahr
- nie

Bemerkung: _____

Welche Ebenen des UP5 nutzen Sie hauptsächlich? (Mehrfachantwort möglich)

- Situation
- Parzellennetz
- Parzellennummern

Bemerkung: _____

Wie nutzen Sie den UP5? (Mehrfachantwort möglich)

- Als alleinige Darstellung
- Als Kartenhintergrund zur Überlagerung mit weiteren Themen

The interface includes a menu bar (File, Edit, View, Insert, Format, Styles, Table, Tools, Window, Help), a toolbar with various icons, and a status bar at the bottom showing "Page 4 of 9", "1,014 words, 8,524 characters", "Default Style", "German (Switzerland)", "Section1", and "100%". A watermark "Activate Windows" is visible in the bottom right corner.

Form controls

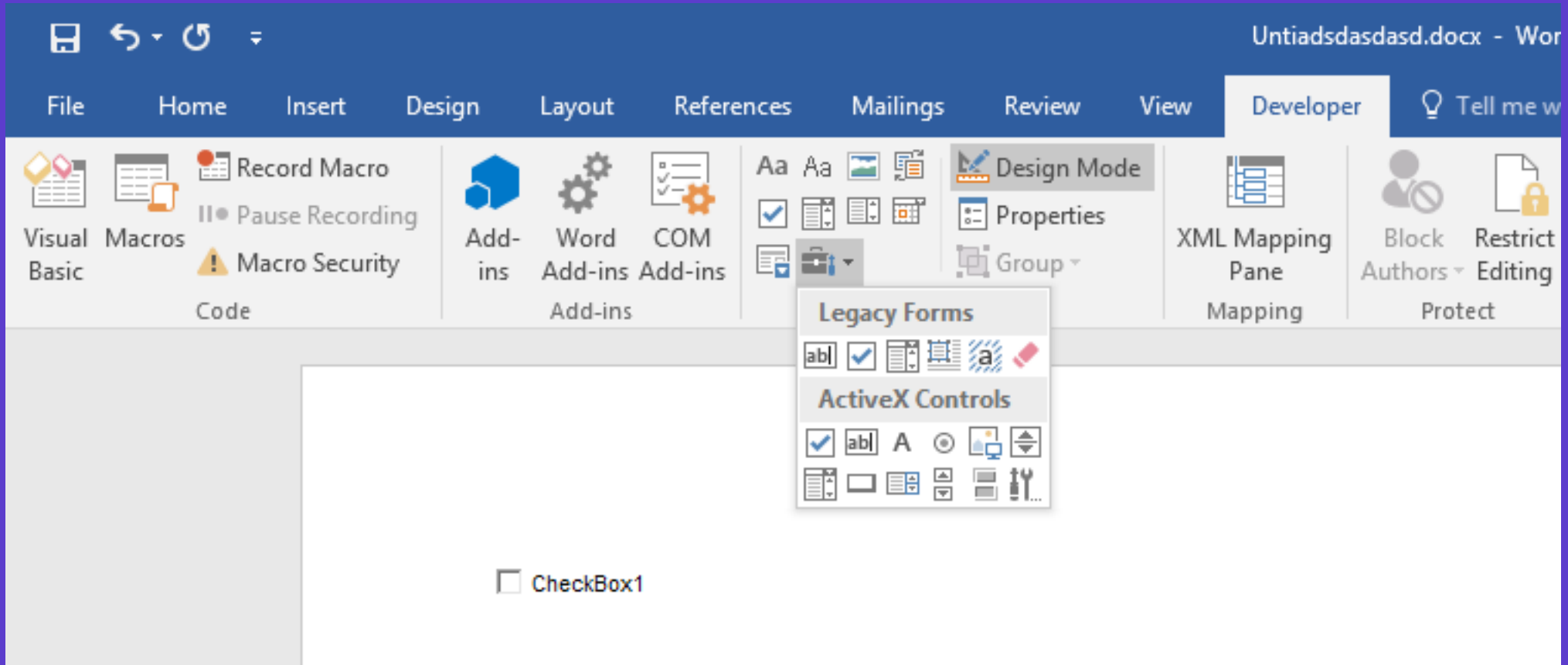
LibreOffice

- Lots of different control types: option button, radio button, pushbutton, combobox, listbox, picture button, etc
- Floating objects: can be anchored as text or as a shape object

MS Office

- Three kind of controls: legacy, ActiveX and content
- Content control: new controls, control types: plain text, image, checkbox, combobox, listbox, date picker
- ActiveX controls: floating, more visual properties (background color, triplstate, wordwrap, image, alignment)

MSO control types



What's done about form controls?

ActiveX controls

- Now imported from DOCX files (tdf#91384)
 - Before only XSLX and PPTX was supported
- LO form controls are exported as MSO ActiveX controls (tdf#50097)
 - ActiveX controls are matching the best to LO form controls
 - Floating objects, similar additional properties, similar types
 - One serious limitation is the missing preset items for list controls
 - Except date picker, which has a better matching with content control



Implementation details

Position, size and anchor

- Specified in /word/content.xml, defined as a VML shape
- A bit different XML tags when the control is a floating object (pict or object context)

Control properties

- Binary file inside DOCX, /word/activeX/activeX1.bin
- /word/activeX/activeX1.xml contains some options how to read that binary
- This binary data seems to be picked up from the DOC binary format

Future possibilities

Check specific controls whether ActiveX the best match

- Maybe list controls (combobox and listbox) would be better to save as content control (because of preset items)

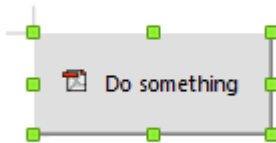
FICHA DE INSCRIÇÃO

Curso	Preliminar de Formação	Data
Dados Pe	Preliminar de Formação	
Nome	Monográfico Manobras de Cordas	
	Monográfico Orientação	
	Monográfico Campo Base	
	Monográfico Animação Escotista	

Future possibilities

Small issues in ActiveX control support (easy hacks)

- ActiveX control locked / read only property is not imported (tdf#111417)
- Additional picture of button control is not saved to DOCX



- Text color of ActiveX controls are not saved to DOCX



Vertical text in a table

Vertical text in MSO

The screenshot displays the Microsoft Word interface with the 'Table Tools' ribbon active. A table is inserted into the document with the following content:

Some text	Some text	Some text

The 'Text Direction - Table Cell' dialog box is open, showing the 'Orientation' section with 'Vertical' selected. The 'Preview' section shows the text 'The quick brown fox jumps over the lazy dog.' oriented vertically. The 'Apply to' dropdown is set to 'Selected cells in the table'. The 'Format Text Effects' task pane is visible on the right side of the window.

Vertical text in a table

MS Office

- Vertical text for most of the container objects (tables, shapes, text boxes)
- Two supported rotation angle (90° and 270°)

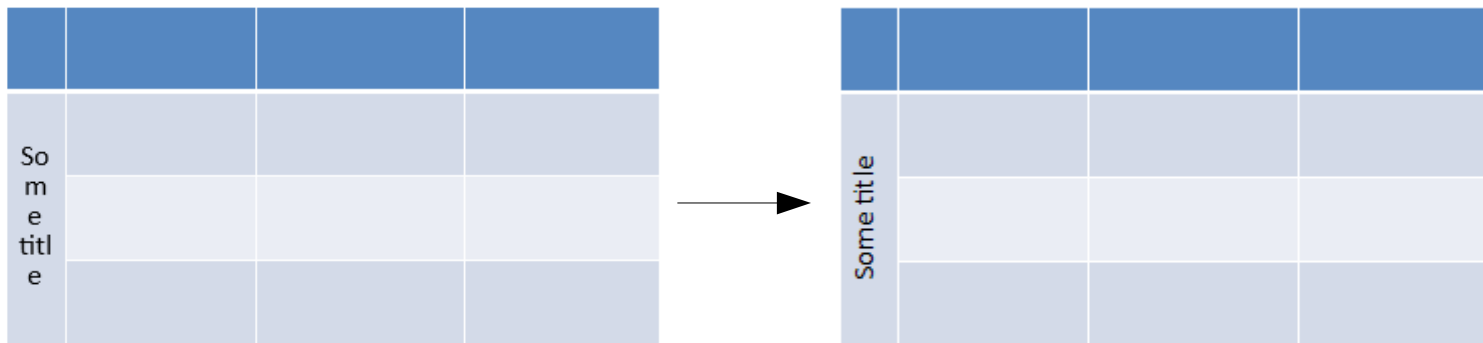
LibreOffice

- Has some vertical text support, but this is related to different language scripts (right-to-left, top-to-bottom, etc)
- Needed to have an actual rotation property which is independent of other font script properties.

What's done about vertical text?

Implemented text rotation for Impress tables

- Now text rotation is imported from PPTX files (tdf#100926)
- The same two rotation angles are supported
- By now it's only available via UNO API, can't set on the GUI



Implementation details

Different levels of text rotation implementation

- Low level code was already there (Calc has text rotation in cells)
- Medium level was added during implementation.
 - Editengine now handles TopToBottom and BottomToTop text.
 - Both during editing and in non-active mode.
- Higher level
 - PPTX import filter now handles the vertical properties (vert, vert270)
 - Calc table cells has this property already, so we did not need to extend API or ODF.

Future possibilities

GUI option for vertical text in Impress tables

- tdf#111633: Missing GUI option for text rotation inside Impress table

Extend vertical text support for other objects

- Writer tables
 - They have a different implementation (sw core)
 - ODF property is already there
- Shapes, text boxes, text frames

Pivot tables

A pivot table example

The screenshot shows a LibreOffice Calc spreadsheet with a pivot table. The pivot table is located in the range C4:D20. The data source is a table with columns for Year (ANNO), Month (MESE), and two Sum of kg prod (I and T). The pivot table is set to show data for the year 2014, with rows for each month and a total row. The pivot table is currently set to show data by month, with the 'pivot' filter selected in the bottom right corner of the pivot table area.

ANNO	MESE	Sum - kg prod I	Sum - kg prod T
2014		1000000	890000
	GENNAIO	250000	200000
	FEBBRAIO	360000	240000
	MARZO	390000	450000
	APRILE		
	MAGGIO	0	0
	GIUGNO		
	LUGLIO		
	AGOSTO		
	SETTEMBRE		
	OTTOBRE		
	NOVEMBRE		
	DICEMBRE		
(empty)	(empty)		
Total Result		1000000	890000

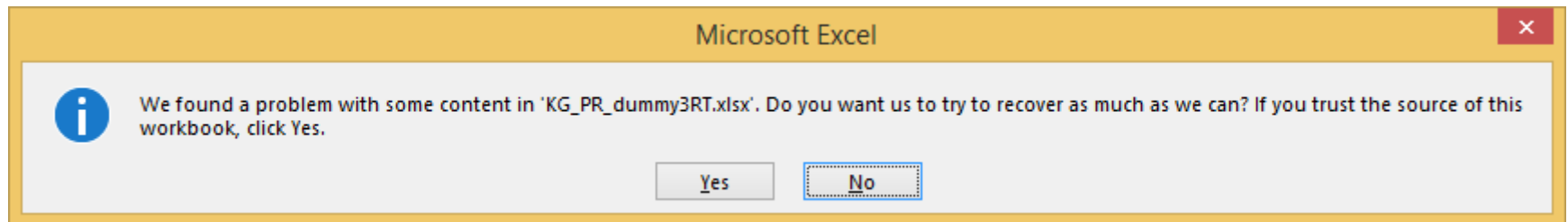
Pivot table

In general

- Special table to organize and aggregate data
- Complex feature supported both by LO and MSO

Issues

- Lots of issues related to MSO compatibility
- Most serious type: corruption and data loss after RT in Calc



What's done about pivot tables?

Performance improvement (tdf#102694)

- Pivot cache behind the pivot table
- Groups are shared between tables
- Specific test document import time 20 min → 0,5 min

Larger and smaller fixes (still working on that)

- Pivot table's layout was wrongly exported (tdf#112946)
- Corrupted XLSX was generated by LO (tdf#112893)
- Filtering of different fields were not preserved (tdf#107711)

Implementation details

MSO pivot table

- Different table layout modes: outline, tabular and compact.
 - LO supports outline and tabular form, but not the compact mode
- Pivot table definition and pivot cache
 - Pivot table: fields, layout, etc
 - Cache contains the shared items, groups, etc.
 - Redundancy in pivot cache definition for robustness (repair dialog)

Future possibilities

Support compact mode in LO

- It needs to have a different pivot table output, with some additional buttons
- LO already support outline and tabular mode, so might be easy to add a new mode (e.g. ODF extension)

Improve performance of XLSX import

- UNO API usage seems slow
- Also pivot cache handling might has issues

Outline mode vs compact mode

KG_PR_dummy (2)asdasdasdas.xlsx - Read-Only - Excel

File Home Insert Page Layout Formulas Data Review View Developer Add-ins Team Tell me what you want to do

Clipboard Font Alignment Number Styles Cells Editing

F23

ANNO	MESE	Somma di kg prod I	Somma di kg prod T
2014		1 000 000	890 000
	GENNAIO	250 000	200 000
	FEBBRAIO	360 000	240 000
	MARZO	390 000	450 000
	APRILE	0	0
	MAGGIO	0	0
	GIUGNO	0	0
	LUGLIO	0	0
	AGOSTO	0	0
	SETTEMBRE	0	0
	OTTOBRE	0	0
	NOVEMBRE	0	0
	DICEMBRE	0	0
Grand Total		1 000 000	890 000

Row Labels	Somma di kg prod I	Somma di kg prod T	
2014	1 000 000	890 000	
GENNAIO	250 000	200 000	
FEBBRAIO	360 000	240 000	
MARZO	390 000	450 000	
APRILE	0	0	
MAGGIO	0	0	
GIUGNO	0	0	
LUGLIO	0	0	
AGOSTO	0	0	
SETTEMBRE	0	0	
OTTOBRE	0	0	
NOVEMBRE	0	0	
DICEMBRE	0	0	
Grand Total		1 000 000	890 000

Activate Windows
Go to PC settings to activate Windows.

Ready 100%



Collabora Productivity

Thank you!

By Tamás Zolnai

tamas.zolnai@collabora.com