OVERVIEW OF LIBREOFFICE 7.0 INTEROPERABILITY FIXES

openSUSE + LibreOffice Virtual Conference, 17.10.2020
Who is speaking?

Vasily Melenchuk

- CIB software GmbH employee
- The Document Foundation member

Michael Stahl

- Joined the Sun Hamburg OpenOffice.org Writer team in 2007
- employed by CIB labs GmbH
Interoperability and OOXXML formats

- OOXXML formats (DOCX, XLSX, PPTX) are popular (obvious!)
- Still in wide use older binary formats and RTF
- LO should import/export them without losing quality
Working on interop bugs

1) Minimize/reduce testcase
2) ?????????
3) Profit!
Working on interop bugs

1) Minimize/reduce testcase
2) ????????
3) Profit! Verify original testcase is fixed
4) Profit, probably
Reducing testcases (steps)

10 Make a file backup
20 Remove something looking unnecessary
30 Verify file is still okay in MS Word
40 Verify file is still broken in Writer
50 GOTO 10
Reducing testcases (details)

- Preferably should be done on XML level, not in Writer or Word
- Need to edit unformatted XML inside zip-archive
  - easier for vim/emacs users
  - sad for others mostly
XML pretty-printing

- In Options->Advanced->Open Expert Configuration find and enable "Pretty Printing"
  - Works only for ODF files
  - Only on save
- "xmllint --format"
- online XML formatters
XML pretty-printing (my solution)

• Primitive script:
  − unpack document package to temp folder
  − xmlint them all
  − pack files back

• https://github.com/bzzzil/doctools
Lists in ODF

L1

L2

Stylesheet

<text:list-style style:name="L1"> ....

<text:list-style style:name="L2"> ....
Lists in DOCX

NumId=1

List Override Table

NumId=1 -> ListId=1
NumId=2 -> ListId=1

List Table

ListId=1 definition
tdf#116883: problem
tdf#116883: format string

- Word: has list level format string "%1-%2-%3-" but it could be even "Hello%1!---%2World!%3 :)
- Writer: only prefix & suffix and hardcoded dot (".") as a separator
- Right now LO uses internally format string
- Not in UI
- prefix/suffix as a fallback an for ODF
- ODF standard need to be extended?
tdf#120394: problem
Lists can refer each other not only via override table, but also with w:styleLink
DOCX lists support: resume

• in total 16 commits with "list" in first commit line (i’m too lazy to verify commits)
• many of them own regressions :( 
• it is not one step forward two step back 
• thanks to Xisco Fauli for assistance & patience :)
Paragraph Mark Formatting

- Word can apply formatting to paragraph mark
  - is also applied to list label
- Writer: there is no paragraph mark
- Status quo: was implemented as a 0 length text attribute at end of the paragraph
  - imported from DOCX paragraph mark→exported to DOCX again
  - a hack that couldn't handle all cases, e.g., if the format of the last character happened to be the same as the paragraph mark
Paragraph Mark Formatting Improvement

• added paragraph item RES_PARATR_LIST_AUTOFMT
• paint as list label
• round-trip DOCX
• hard formatting in 6.4.0
• character style too in 6.4.7
• TODO: RTF, ODF extension, UI, paint as pilcrow too
Example
CIB
IDEEN MIT SYSTEM

OUR PRODUCTS:
https://libreoffice.cib.de