



RTF támogatás a LibreOffice Writer programban

- ▼ Vajna Miklós
- ▼ 2011. november 12.



All text and image content in this document is licensed under the [Creative Commons Attribution-Share Alike 3.0 License](#) (unless otherwise specified). "LibreOffice" and "The Document Foundation" are registered trademarks. Their respective logos and icons are subject to international copyright laws. The use of these therefore is subject to the [trademark policy](#).

Bevezető

- ▼ A Google Summer of Code-ról
- ▼ A LibreOffice fejlesztésről
- ▼ RTF export fejlesztés
- ▼ RTF import fejlesztés

Google Summer of Code

- ▼ Pályázati lehetőség szervezeteknek, tanulóknak
- ▼ Szervezetek projektötletekkel pályáznak
- ▼ Tanulók (ötletek alapján) kidolgozott projekttervvel
- ▼ Szabad szoftverek fejlesztése
 - ▼ Tanulók pénzt kapnak: Flip Bits not Burgers

LibreOffice fejlesztés

- ▼ 2010. szeptember 28-án indult a projekt
- ▼ The Document Foundation, mint jogi háttér
- ▼ OpenOffice.org fork, első lépésben a Go-OO átnevezve
- ▼ Az ottani tudás nagyrészt alkalmazható itt is
- ▼ A kód nagy része C++ (mítosz: Java)
 - ▼ De: XML, Make, ASM, Yacc, Perl, Python, stb.

LibreOffice első build

▼ Forrás letöltése:

```
git clone git://anongit.freedesktop.org/libreoffice/core
```

▼ Fordítás:

```
./autogen.sh  
make  
make dev-install
```

▼ Futtatás:

```
cd install/program  
source ./ooenv  
./soffice.bin
```

Inkrementális build

- ▼ Egy teljes build sok idő (ld. később)
- ▼ Fejlesztés közben elengedhetetlen a gyors iteráció
- ▼ Egy-egy modult külön-külön is újra lehet fordítani:

```
writerfilter$ time make -sr -j4 dbglevel=2 build
[ build CXX ] writerfilter/source/rtftok/rtfdocumentimpl
[ build DEP ] LNK:Library/librtftoklo.so
[ build LNK ] Library/librtftoklo.so
[ build LNK ] Library/libwriterfilterlo.so
[ build MOD ] writerfilter
[ build ALL ] top level modules: writerfilter
[ build ALL ] loaded modules: writerfilter
real      0m6.170s
```

Projekt méret

- ▼ Egyik legnagyobb free software projekt
 - ▼ split build nem jött be
- ▼ ~8 millió kódsor
- ▼ Forráskód mérete ~4GB
- ▼ Fordítási idők szélsőértékei Linuxon:
 - ▼ Fejlesztői build (csak angol, help nélkül, stb.):

```
runtime: 15.30 (minutes)
```

- ▼ Minden nyelvvel:

```
Making: libreoffice 3.4.3.2-3 (Wed Oct 19 08:23:16 CEST 2011)  
Finished: libreoffice (Wed Oct 19 11:58:08 CEST 2011)
```

Kollaboráció

- ▼ Gitet használunk: elosztott verziókezelés
- ▼ Választás szabadsága:
 - ▼ Patch URL-jének linkelése IRC-n
 - ▼ Patch küldése a fejlesztői listára
 - ▼ Saját branch-be pusholni
 - ▼ Saját repóba
- ▼ Mára egyszerűbb: legtöbb változtatás egyetlen repóba
- ▼ Stratégiai döntések: Technical Steering Call

LibreOffice filterek

- ▼ Model-View-Controller: a filterek a modellt mentik el / töltik be
- ▼ Export filter: egy létező modellt szerializál egy szabványnak megfelelő formátumba (ideális esetben)
- ▼ Import filter: egy üres modellt kap, abba építi fel az input alapján a dokumentumot
- ▼ Ez után nem meglepő, ha egyetlen karaktert átírva a dokumentumban mentés után a kimenet gyökeresen eltér...
- ▼ Egy-egy filter egy LibreOffice alkalmazáshoz készül, továbbiakban csak Writer filterekkel foglalkozunk

Miért izgalmas az RTF formátum?

- ▼ Első verzió: 1987 (XML: 1998, ODF: 1999)
- ▼ Újabb verziói visszafele kompatibilisak
- ▼ Időről időre frissül, a 1.9.1-es verzió képes eltárolni minden fontosabb mai funkciót is (beágyazott táblázatok, OLE objektumok, stb.)
- ▼ Előszeretettel használják sok helyen, ahol az ODF még túl új (Microsoft világban az Office 2010 az első, ami natívan képes kezelni.)

RTF export fejlesztés

- ▼ 2010-ben egy teljesen új RTF export filter készült el
- ▼ Ötlet: az RTF sokban hasonlít a doc/docx-re
 - ▼ Microsoft találmány
- ▼ A doc/docx exporternek van már egy közös alapja
- ▼ Készítsünk egy új RTF exportert erre az alapra építve
- ▼ Cél:
 - ▼ Tudjon mindent amit a régi
 - ▼ Kisebb méret
 - ▼ Új funkciók

RTF export integráció

- ▼ Először a Go-OO 3.3-ban jelent meg
- ▼ Része lett az OpenOffice.org 3.4 bétának
 - ▼ Amiből azóta sincs stabil kiadás
 - ▼ Ld. Apache szappanopera
 - ▼ Ettől függetlenül: az Oracle QA hasznos volt
- ▼ Tipikus use case: ODF dokumentum RTF-be mentve
 - ▼ Majd eredmény ellenőrzése Microsoft Word alatt

Új export funkciók: könyvjelzők

▼ Régi:

```
Hello world! 1¶  
reference to footnote: Error: Reference source not found¶
```

▼ Új:

```
Hello world! 1¶  
reference to footnote: 1¶
```

Új export funkciók: karakter tulajdonságok

▼ Régi:

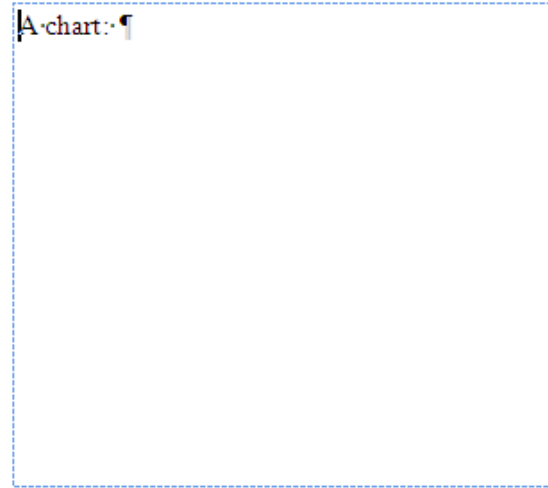
```
expanded·space·by·5·and·10pt¶  
VA·no·kerning:·VA¶
```

▼ Új:

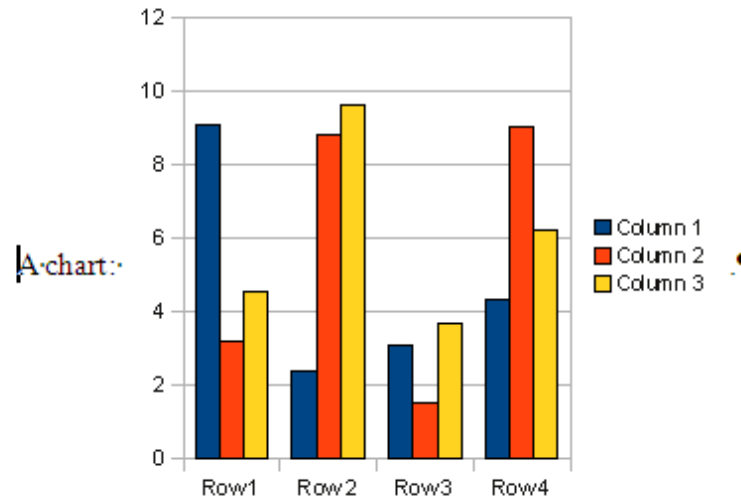
```
expanded·s p a c e ·by·5·and·10pt¶  
VA·no·kerning:·VA¶
```

Új export funkciók: diagram, egyéb OLE objektum

▼ Régi:



▼ Új:






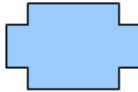

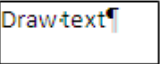
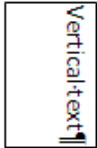
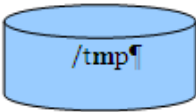


Új export funkciók: rajzok

▼ Régi:

```
Rectangle.¶  
¶  
Ellipse.¶  
¶  
¶  
Freeform-line.¶  
¶  
Line.¶  
¶  
Basic-shape-examples.¶  
→ regular-pentagon,-cross,-cube¶  
¶  
Text.¶  
¶  
Vertical-text.¶  
¶  
Magnetic-disc-with-formatted-text.¶
```

▼ Új:

```
Rectangle.¶   
¶  
Ellipse.¶   
¶  
Freeform-line.¶   
¶  
Line.¶   
¶  
Basic-shape-examples.¶     
→ regular-pentagon,-cross,-cube¶  
¶  
Text.¶   
¶  
Vertical-text.¶   
¶  
Magnetic-disc-with-formatted-text.¶ 
```

Új export funkciók: űrlapok

▼ Régi:

```
This one is selected:   
This is a textbox:   
Listbox: 

- 

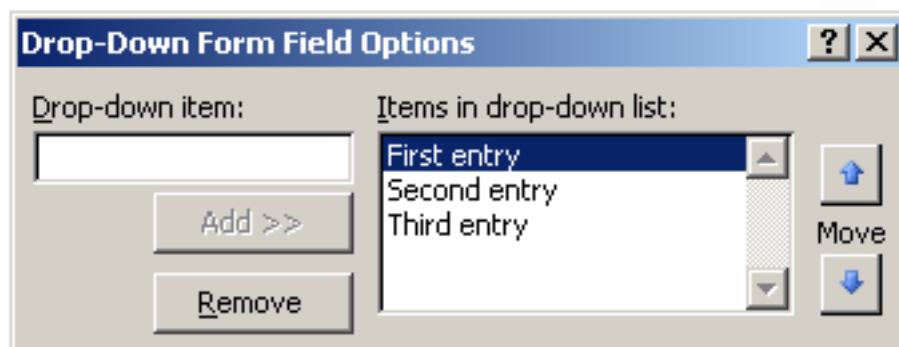
```

▼ Új:

```
This one is selected:   
This is a textbox:   
Listbox: 

- Third entry

```



Új export funkciók: sorszámozás

▼ Régi:

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis rhoncus imperdiet velit, sed luctus augue pellentesque a. Cras semper purus eu est porta euismod. Sed auctor laoreet erat, pretium condimentum libero condimentum sit amet. Aenean lectus purus, blandit ut volutpat eget, consequat eu leo. In dignissim rhoncus libero ut sollicitudin. Donec egestas viverra mi eu consequat. Maecenas malesuada sagittis posuere. Curabitur facilisis, dui nec cursus viverra, felis arcu vulputate magna, at fringilla est neque pretium ipsum. Duis in velit diam, bibendum vehicula diam. Cras nunc ligula, condimentum non tincidunt et, interdum sit amet enim. Sed ac felis vitae justo tempus dignissim. Praesent dui massa, imperdiet id dapibus at, egestas sed enim. Nam bibendum blandit nisi ac dignissim. Cras eget urna justo. Sed mi urna, facilisis et mollis a, rutrum ac libero. Donec pellentesque luctus posuere.¶

▼ Új:

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis rhoncus imperdiet velit, sed luctus augue pellentesque a. Cras semper purus eu est porta euismod. Sed auctor laoreet erat, pretium condimentum libero condimentum sit amet. Aenean lectus purus, blandit ut volutpat eget, consequat eu leo. In dignissim rhoncus libero ut sollicitudin. Donec egestas viverra mi eu consequat.

5 Maecenas malesuada sagittis posuere. Curabitur facilisis, dui nec cursus viverra, felis arcu vulputate magna, at fringilla est neque pretium ipsum. Duis in velit diam, bibendum vehicula diam. Cras nunc ligula, condimentum non tincidunt et, interdum sit amet enim. Sed ac felis vitae justo tempus dignissim. Praesent dui massa, imperdiet id dapibus at, egestas sed enim. Nam bibendum blandit nisi ac dignissim. Cras eget urna justo. Sed mi urna, facilisis et mollis a, rutrum ac libero. Donec

10 pellentesque luctus posuere.¶

Új export funkciók: matematikai kifejezések

▼ Régi:

before

▼ Új:

before $a = \frac{b}{c}$ after



Új export funkciók: oldaltörések

- ▼ Teszt: „sima” oldaltörés, majd jobb és megint jobb
 - ▼ Balról indulva ez összesen 5 oldal kell legyen
- ▼ Régi:

Then a „right page” break after an odd page, should lead us to page #5.¶

Page: 3 of 4 Words: 185 Hungarian

- ▼ Új:

Then a „right page” break after an odd page, should lead us to page #5.¶

Page: 5 of 5 Words: 185 Hungarian

Új export funkciók: Oldalszámozás

▼ Régi:

→ 1¶
This is the first page with roman numerals. ¶Section Break (Next Page)
→ 2¶
This is the first page with „Default” style, with page numbering restarted. ¶

▼ Új:

▪ → i¶
This is the first page with roman numerals. ¶Section Break (Next Page)
▪ → 1¶
This is the first page with „Default” style, with page numbering restarted. ¶

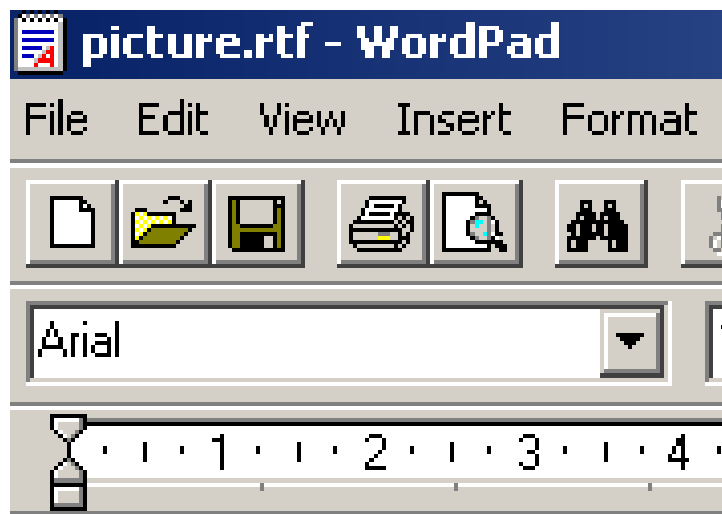
Új export funkciók: Képek Wordpadben

▼ Régi:



Twitter!

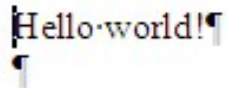
▼ Új:



 Twitter!

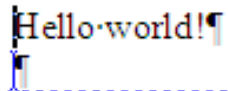
Új export funkciók: Postit mező

▼ Régi:

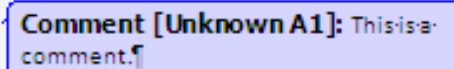


Hello world!

▼ Új:



Hello world!



Comment [Unknown A1]: This is a comment.

Új export funkciók: Szakaszok / hasábtörés

▼ Régi:

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Column break. ¶
 Vivamus elit tellus, accumsan in congue eget, ullamcorper id odio. Ut nunc est, rutrum non facilisis ac, iaculis sit amet mauris. Curabitur ut magna ut odio vehicula scelerisque vitae at ante. Aliquam erat volutpat. Vivamus varius augue ut metus tristique id viverra nulla semper. Curabitur volutpat est tortor, sed dictum ipsum. Nam sed neque ante, nec sodales lorem. Quisque eleifend auctor pulvinar. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Duis et tempus velit. Donec porttitor nisi et est feugiat vulputate sed eget urna. Fusce facilisis, mi a suscipit venenatis, turpis lacus egestas odio, in tempus quam lacus ut est. ¶

▼ Új:

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Column break. ¶

Vivamus elit tellus, accumsan in congue eget, ullamcorper id odio. Ut nunc est, rutrum non facilisis ac, iaculis sit amet mauris. Curabitur ut magna ut odio vehicula scelerisque vitae at ante. Aliquam erat volutpat. Vivamus varius augue ut metus tristique id viverra nulla semper. Curabitur volutpat est tortor, sed dictum ipsum. Nam sed neque ante, nec sodales lorem. Quisque eleifend auctor pulvinar. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae; Duis et tempus velit. Donec porttitor nisi et est feugiat vulputate sed eget urna. Fusce facilisis, mi a suscipit venenatis, turpis lacus egestas odio, in tempus quam lacus ut est. ¶

Új export funkciók: Védett szakaszok

▼ Régi:

```
¶.....: Section Break (Continuous).....  
This section is not write-protected (without a password). ¶.....: Section Break (Continuous).....  
¶
```

▼ Új:

```
¶.....: Section Break (Continuous).....  
This section is write-protected (without a password). ¶.....: End of Protected Section.....  
¶
```

Új export funkciók: Beágyazott táblázatok

▼ Régi:

Nested table:

A		B
Nested-a	Nested-b	D
Nested-c	Nested-d	
	D	
A row after nested table		F

▼ Új:

Nested table:

A		B
Nested-a	Nested-b	D
Nested-c	Nested-d	
A row after nested table		F

Új export funkciók: Tartalomjegyzék

▼ Régi:

▪ TOC¶	Section Break (Continuous).....	
▪ Table of Contents¶		
First heading¶	→	1¶
Second heading¶	→	1¶
Third heading¶	→	1¶
TOC¶	→	1¶
¶		

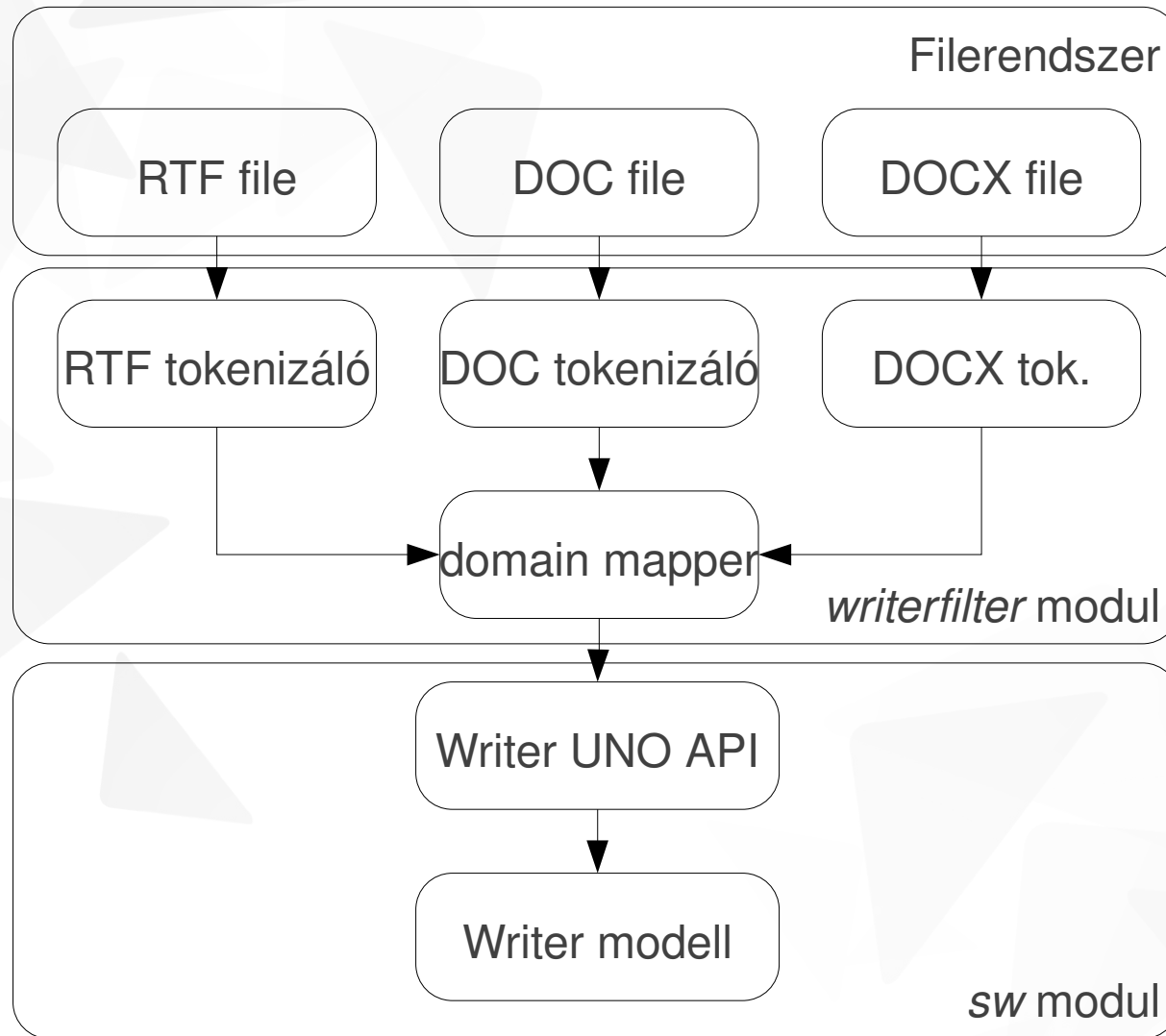
▼ Új:

→ TOC¶	Section Break (Continuous).....	
▪ Table of Contents¶	Section Break (Continuous).....	
First heading	→	1¶
Second heading	→	1¶
Third heading	→	1¶
TOC	→	1¶
¶		

RTF Import fejlesztés

- ▼ Ötlet itt is hasonló
 - ▼ Használjuk a docx importer frameworkjét
 - ▼ A writerfilter modul már tartalmaz egy domain mappert
 - ▼ A Writer vs. Word problémák már itt meg vannak oldva
 - ▼ Például mezők értelmezése
 - ▼ A célok változatlanok
 - ▼ Regressziók elkerülése
 - ▼ Kisebb méret, új funkciók

RTF import architektúra



RTF import tesztelés

- ▼ Unit teszt: csak a tokenizert teszteli
 - ▼ Gyors: 215ms az összes teszt
- ▼ Már rögtön a writerfilter végén lefut
 - ▼ Ekkor az sw modul még akár le se fordult
- ▼ Kézi tesztelést nem helyettesíti
 - ▼ Vizuális egyezést nem vizsgál
- ▼ Tesztelt dokumentumok:
 - ▼ Korábbi CVE hibák (unit teszt)
 - ▼ Word 2007, 2010, OOo 3.3, LO 3.4 RTF kimenetek

Új RTF import funkciók: beágyazott táblázatok

▼ Régi:

```
Nested-table:¶  
A¶  
B¶  
¶  
C¶  
D¶  
¶  
C'¶ B'D'¶  
¶
```

C'		B'D'
----	--	------

▼ Új:

```
Nested-table:¶  
A¶ B¶ B'¶  
C¶ D¶  
C'¶ D'¶  
¶
```

A	B	B'
C	D	
C'	D'	

Új RTF import funkciók: lábjegyzetek

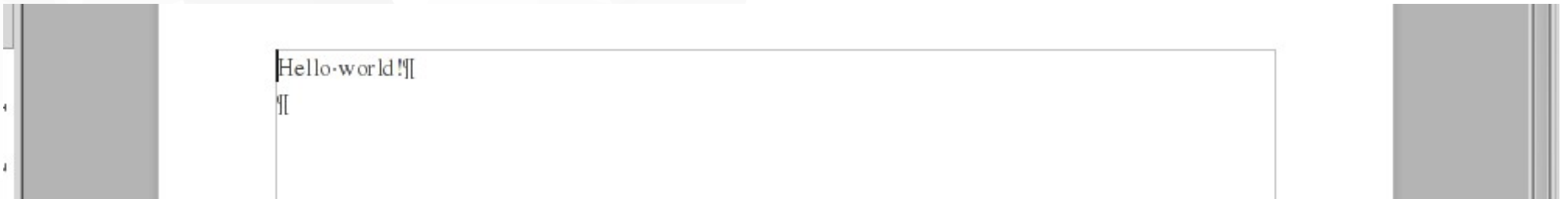
- ▼ Régi:
 - ▼ Hiányzó felső index
 - ▼ Több karakter esetén csak az utolsó a mező része
- ▼ Új:

```
Hello·world!¹¶  
Foobar.*¶  
Some·text.i¶  
More.**¶
```

```
Hello·world!¹¶  
Foobar.*¶  
Some·text.i¶  
More.**¶
```

Új RTF import funkciók: post-it mezők

▼ Régi:



▼ Új:



Új RTF import funkciók: űrlapok

▼ Régi:

```
This is a checkbox with help text:>[]  
This one is selected:>[]  
This is a textbox: Current text>[]  
Listbox:>[]
```

▼ Új:


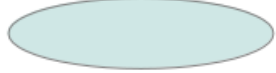

```
This is a checkbox with help text: ¶  
This one is selected: ¶  
This is a textbox: Current text¶  
Listbox: Third entry¶
```

Új RTF import funkciók: rajzok

▼ Régi:

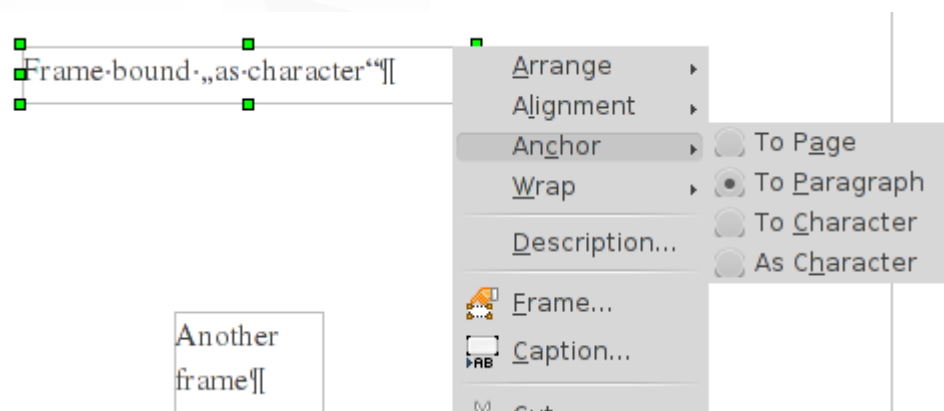
```
Rectangle:¶  
¶  
Ellipse:¶  
¶  
¶  
  
Freeform line:¶  
¶  
¶
```

▼ Új:

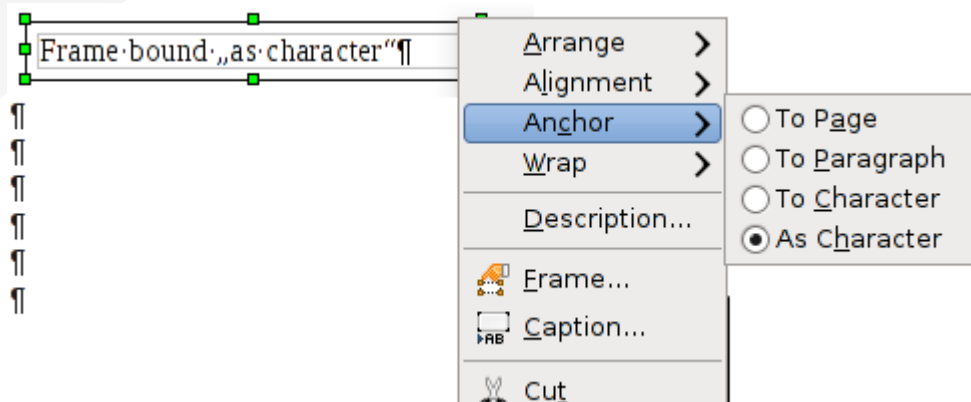
```
Rectangle:¶   
¶  
Ellipse:¶   
¶  
¶  
  
Freeform line:¶   
¶  
¶
```

Új RTF import funkciók: szövegdobozok

▼ Régi:



▼ Új:



DOCX import mellékhatások

▼ Szöveg-szöveg igazítás, dupla áthúzás

▼ Régi:

text-to-text alignment: top

single strike-through (there are 4 other types but only double is supported by rtf)

superscript

subscript

expanded spacing

rotation with 90

before text of the hyperlink after

▼ Új:

text-to-text alignment: top

single strike-through (there are 4 other types but only double is supported by rtf)

superscript

subscript

expanded spacing

rotation with 90

rotation

before text of the hyperlink after

DOCX import mellékhatások

▼ Lábjegyzetek újraszámozása, extra bekezdés

▼ Régi:

ctetur adipisc ing elit. Maecenas condimentum consectetur est,
t. Suspendisse potenti. In hac habitasse platea dictumst. Praesent
cus voluptat. In vitae arcu sit amet velit commodo placerat in ut
ondimentum iaculis. Maecenas ac justo mauris, sit amet
s eros. Suspendisse cursus, nulla in faucibus dictum, purus lorem
it ut tortor. Suspendisse potenti. Curabitur luctus malesuada nisl,
z ullamcorper porta cursus. Quisque consequat semper nunc, eu
que mauris neque, facilisis sed molestie at, ultrices sed enim.
at. ¶

Vestibulum dapibus tincidunt cor
amet consectetur mauris. Quisqu
turpis nec lectus rutrum rhoncus
ultrices consectetur varius. Cras
Nam vitae tortor a ligula interdum
vehicula nibh adipisc ing. ¶

1 Lorem ipsum dolor sit amet, consectetur adipiscing elit. ¶

▼ Új:

tetur adipisc ing elit. Mae cenas condimentum consectetur est,
t. Suspendisse potenti. In hac habitasse platea dictumst. Praesent
is voluptat. In vitae arcu sit amet velit commodo placerat in ut
ndimentum iaculis. Maecenas ac justo mauris, sit amet
eros. Suspendisse cursus, nulla in faucibus dictum, purus lorem
ut tortor. Suspendisse potenti. Curabitur luctus malesuada nisl,
ullamcorper porta cursus. Quisque consequat semper nunc, eu
ue mauris neque, facilisis sed molestie at, ultrices sed enim.
: ¶

Vestibulum dapibus tincidunt cor
amet consectetur mauris. Quisque
turpis nec lectus rutrum rhoncus u
ultrices consectetur varius. Cras a
vitae tortor a ligula interdum aliq
vehicula nibh adipisc ing. ¶

1 Lorem ipsum dolor sit amet, consectetur adipiscing elit. ¶

Hivatkozások

- ▼ LibreOffice: <http://www.libreoffice.org/>
- ▼ GSoC: <http://code.google.com/soc/>
- ▼ Új RTF export:
<http://cgit.freedesktop.org/libreoffice/core/tree/sw/source/filter/ww8>
- ▼ Új RTF import:
<http://cgit.freedesktop.org/libreoffice/core/tree/writerfilter/source/rtftok>
- ▼ Diák: <http://vmiklos.hu/odp/>



Köszönöm a figyelmet ...

▼ ... kérdések?



All text and image content in this document is licensed under the [Creative Commons Attribution-Share Alike 3.0 License](#) (unless otherwise specified). "LibreOffice" and "The Document Foundation" are registered trademarks. Their respective logos and icons are subject to international copyright laws. The use of these therefore is subject to the [trademark policy](#).