

Reflektorfényben az OpenOffice.org

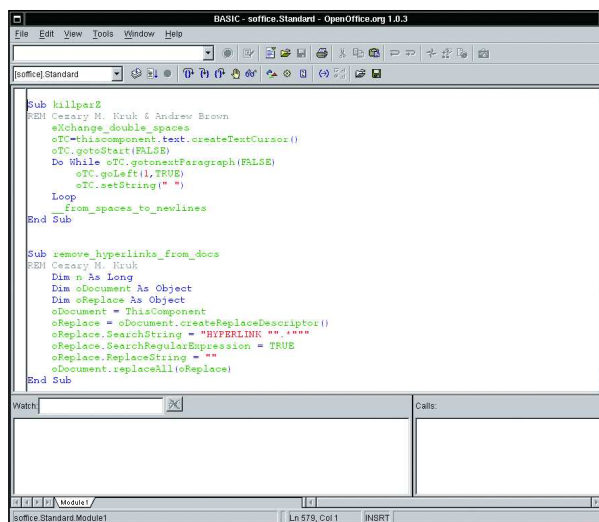
Vajon jól szolgál az OpenOffice.org Writer, mint szerkesztői eszköz egy többféle géptípust is használó kiadóvállalatnál?

Az *OpenOffice.org* egy nagyszerű alkalmazáscsomag, számos hasznos összetevőből áll, melyek számtalan lehetőséget kínálnak használójuknak. Testreszabható, és jó néhány nyílt dokumentumformátumot támogat. Ha az alapbeállításokat saját, egyéni igényeinkhez szeretnénk igazítani, az *OpenOffice.org* makrók és egyéb parancsfájlok készítésének lehetőségével van segítségünkre. Szerkesztőként dolgozom egy szabad szoftverekkel foglalkozó lencyel magazinnál. A szerkesztési folyamat kezdetén a szerző átadja a szöveget a szerkesztőnek, aki átszerkeszti azt. Az átszerkesztés a különféle tartalmi és formai hibák javítását jelenti, illetve a szöveg felkészítését, szabványos formába hozatalát a további lépésekben történő feldolgozáshoz. Ezután a korrektor feladata a szöveg további javítása, majd újra a szerkesztő lép a színre, aki elvégzi a végső módosításokat. Végül a szedő előkészíti a szöveget a nyomtatásra, majd a teljes anyag utolsó ellenőrzését újfent a szerkesztő végzi el. A szöveg minden egyes lépésnél más-más formát ölt. A mi kiadóvállalatunk a nyílt dokumentumformátumokat támogatja, ezért szerzőink *HTML* formátumban vagy egyszerű szöveggé szálítják anyagaikat, a képeket pedig *PNG* vagy *EPS* formátumban mellékelik. Az anyag szerkesztése után a szerkesztő küld egy másolatot a szerzőnek, ez *HTML* formátummal történik. A korrektorok *Microsoft Windows* rendszereken dolgoznak, *Microsoft Worddel*, ezért ők .doc formátumban igénylik az írásokat. A szedők Macintosh gépeket és *QuarkXPress*t használnak. Nekik kétféle dokumentumra van szükségük: a *Microsoft Word* fájlra a nyomtatáshoz és a cikkhez előírt formátum ellenőrzéséhez, valamint egy *Macintosh* szövegfájlra, amivel *Quark* alatt is tudnak dolgozni.

Negyedévente megjelenő kiadványunk 2000 őszén indult, akkor még *StarOffice*-ot használtam. Azóta áttértem az *OpenOffice.org*-ra. A szerzők szövegfájljaival *StarOffice* és *OpenOffice.org* alatt nagyon hasonlóan tudok dolgozni. Korábban *StarWriter*, jelenleg *OpenOffice.org Writer* alatt beimportálom a szöveges vagy *HTML* formátumban lévő dokumentumot, majd amikor végeztem vele, *HTML*, *Microsoft Word*, illetve *SDW* vagy *SXW* formátumba mentem.

Szöveg- és HTML fájlok importálása

Ha egy forrásfájl megfelelő formában jut el hozzánk, akkor a megnyitásával nem lehet gond. Ha egy fájl megsérült,



1. ábra A KillparZ makró végzi el a beimportált szövegfájlok előzetes feldolgozását

akkor helyre kell állítani. Figyelembe véve a dokumentumok nyílt formátumát, ez nem jelenthet különösebb kihívást. Miután beimportáltunk egy fájlt, a megfelelő formátumra kell alakítani. A lencyel, német, francia és egyéb nem angol nyelvű írások szerkesztőinek a kódlapot is át kell állítaniuk. A lencyel dokumentumoknál például a szabványos kódlap az *ISO-8859-2*, viszont az összes OpenOffice.org dokumentum alap kódlapja az *UTF-8*. Ha a beimportált dokumentumokat kényelmesen akarjuk átalakítani, akkor egy makróra lesz szükségünk. Az általam *OpenOffice.org* alá készített makrók több kódlap-átalakítót is tartalmaznak, például kérések az *ISO-8859-2* – *UTF-8* és fordított irányú átalakításra. A szövegszerkesztők némelyikében készült szövegfájlok bekezdései sokszor több sorba kerülnek. Összevonásukra a *KillparZ* makrókat használhatjuk, ez az *Andrew Brown* által készített killpars továbbfejlesztett változata (1. ábra). A *KillparZ* az *ooo-macro* csomag egyik eleme. Feltéve, hogy a szerző a megfelelő kódlapot adta meg, a *HTML* fájlok importálásakor sem lehet gond a karakterkészlettel. Az viszont kellemetlen meglepetésként érhet, hogy a *HTML* dokumentumok esetében a makrókhoz rendelt gyorsbillentyűk nem működnek. Ahhoz, hogy szóra



4. ábra A CHIP Special szerkesztői gárdája, balról jobbra: Robert Bielecki (szerkesztő), Romek Gnitecki (főszerkesztő), Cezary M. Kruk (CHIP Special Linux) és Tomek Borukalo (szerkesztő)

Az általunk kezelt írások egyszerű dokumentumok. Szerkesztőségünk a fent említett három betűkészetet használja, természetesen dőlt és félkövér betűket is alkalmazunk, továbbá kétféle fejléct és egyszerű táblázatokat. A dokumentumokba nem szoktunk képeket illeszteni, egyszerűen csak felsoroljuk bennük a *PNG* vagy *EPS* formátumú képfájlok neveit. Ezeket a dokumentumokat *SDW* vagy *SXW* formátumból is gond nélkül lehet Microsoft Word formátumba kimenteni.

HTML formátum

A *HTML* formátumú dokumentumok előállítására már kicsit nehezebb feladat. A *StarWriter* és az *OpenOffice.org Writer* kétségkívül kifinomult *HTML* fájlokat állítanak elő, amint az a 2. ábrán is látható. Semmi nem akadályozza meg ugyanakkor, hogy ezeket a fájlokat egy egyszerű Perl parancsfájllal továbbalakítsuk. Nálam ez a parancsfájl a *soffice2html* nevet kapta. A parancsfájl elején a sorvégeket cseréljük le szóközökre, valahogy így:

```
s/\n/ /;
```

Ezután a kód bizonyos elemeit cseréljük le másokra. Például:

```
s/<(\/?)>B>/<$1STRONG>/g;  
s/<(\/?)>I>/<$1EM>/g;
```

A fenti parancsokkal az összes ` ... ` és `<I> ... </I>` címkepárt lecseréljük ` ... ` és ` ... ` párokra, így a félkövér és a dőlt szakaszok jelölése szabványos lesz. Ezután vegyük ki a felesleges címkeket, például:

```
s/<EM><EM>/<EM>/g;  
s/<\/EM><\/EM>/<\/EM>/g;
```

A sorvégek egy részét érdemes visszaállítani. Használjuk például a következő parancsokat:

```
s/(.+?)</$1\n/g;  
s/>(.+?)>/>\n$1/g;
```

Ezzel az egyes *HTML*-címkék elé és mögé sorvégejeleket helyezünk el. Ha szakemberhez méltó parancsfájlt akarunk írni, talán még egy záró parancsot adjunk hozzá:

```
print OUT "<!-- ", "soffice2html: ",  
        scalar localtime, " ->\n";
```

A záró parancssal egy az alábbihoz hasonló megjegyzés kerül a feldolgozáson átesett *HTML* fájlba:

```
<!-- soffice2html: wed Jul 23 17:34:35 2003 -->
```

Tehát, ha a *dokumentum.sxw*-t kimentjük, mint *dokumentum.html*-t, akkor ez utóbbit a *soffice2html* *dokumentum.html* parancssal tudjuk feldolgozni. (3. ábra) A *HTML* fájlok ilyen jellegű átalakításával érhetjük el a legjobb eredményt, vagyis szabványosabb, könnyebben olvasható kódhoz és 15-40 százalékkal kisebb fájlokhoz jutunk. A *soffice2html* parancsfájl szintén megtalálható az *ooo-macro* csomag jelenlegi változatában.

Ha egy dokumentumból egyszerű *Macintosh* szövegfájl akarunk előállítani, akkor a megfelelő karakterkészletet használó „*Text Encoded*” fájltypussal kell mentenünk. A lengyel dokumentumoknál a megfelelő készlet az *Eastern Europe (Kelet-Európa)*.

Ez a kimentési mód a hétköznapi feladatokhoz tökéletesen megfelel, de nyomdai célokra nem. Cikkeinkben sokszor szerepelnek a különféle műveletekhez tartozó billentyű-kombinációk szimbólumai és egyéb különleges karakterek. Ha a normál eljárást követjük a *Macintosh* szövegfájlok előállításakor, akkor ezeket a karaktereket elveszítjük. Ahhoz, hogy megmaradjanak, egy makróra van szükségünk, amely az *UTF-8* kódlap szerinti karaktereket *Macintosh* kódlap szerintiekre alakítja. Az e célra használható makró, a *recode_utf_8_to_apple_macintosh* szintén az *ooo-macro* csomag része.

A makróval úgy készíthetünk szövegfájlt, hogy lefuttatjuk, majd a dokumentumot „*Text Encoded*” típusal mentjük, a *System* karakterkészlet és *CR* bekezdésjelek használatával. A fájlba így minden a szedő munkáját segítő és megkönnyítő információ bekerül.

Reflektorfényben

Az *OpenOffice.org Writer*, mint szerkesztői eszköz segítségével elvégezhetjük a dokumentumok feldolgozását, megoszthatjuk őket a szerzők, a korrektorok és a szedők között – s mindez az összes résztvevő számára kényelmes, átlátszó módon történik. Mindössze a *Writer*re, néhány *TrueType* betűkészletre és makróra, valamint a letisztultabb *HTML* fájlokat előállító Perl parancsfájllal van szükségünk.

Linux Journal 2005. február, 130. szám



Cezary M. Kruk Lengyelországban, Wrocław városában él. A negyedévente megjelenő lengyel *CHIP Special Linux* egyik szerkesztője.