

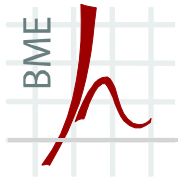
# 13. Fájlformátumok

Schulcz Róbert

[schulcz@hit.bme.hu](mailto:schulcz@hit.bme.hu)

Madarassy László

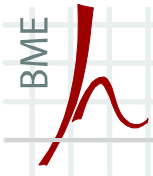
[lmadarassy@mik.bme.hu](mailto:lmadarassy@mik.bme.hu)



# Témakörök

---

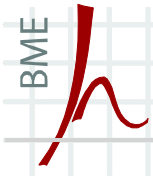
- Microsoft Word verziók (doc, docx, dot, stb.)
- OpenOffice.org (odf, odt)
- Univerzális formátumok (RTF, TXT, PDF)
- Fájlok konvertálása
  - PDF készítés



# Microsoft Word formátumok 1.

---

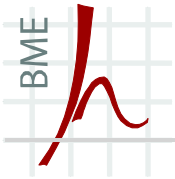
- 1981-ban jelent meg az első változat
- 14 verzió jelent meg azóta (DOS, Windows és Mac OS-re)
- Az újabb változatokkal készült fájlok megnyitása általában nem működik tökéletesen a régebbi változatokon



# Microsoft Word formátumok 2.

---

- Word for DOS
- Word for Windows 1.0, 2.0 és 4.0, 5.0 for Mac
- Word 6 és Word 95 for Windows, Word 6.0 for Mac
- Word 97, 2000, 2002, 2003 for Windows, Word 98, 2001, X, 2004 for Mac
- Word 2007 for Windows, Word 2008 for Mac (docx: Open XML)



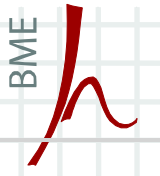
# Word 97-2003 formátuma

---

- Bináris formátum, un. OLE (Object Linking and Embedding) struktúrát használ. Az OLE úgy viselkedik, mint egy merevlemez, 512 byte-os blokkokból épül fel.
- Felépítése:
  - header blokk: tartalmazza a főbb elemek helyét az adattárolóban (mint merevlemeznél a FAT tábla).
  - file location blokk: ki hozta létre, mikor módosították, stb.

# Word 2007 formátuma 1.

- Kiterjesztése: DOCX, XLSX
- Egy ZIP file (Open Packaging Structure), amiben XML fájlok, képek, leírok találhatóak.
- Office Open XML formátum: nem szabványos, az XML szabványt sem követi teljesen.
- WordprocessingML, SpreadsheetML, PresentationML leíró nyelvek használata.
- A meta információk leírása az un. Dublin core-nak megfelelően történik.
- Az alap dokumentum tartalmaz egy [Content\_Types].xml fájlt.



# Word 2007 formátuma 2.

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<Types xmlns="http://schemas.openxmlformats.org/package/2006/content-types">
  <Default Extension="png" ContentType="image/png"/>
  <Default Extension="emf" ContentType="image/x-emf"/>
  <Default Extension="rels" ContentType="application/vnd.openxmlformats-
    package.relationships+xml"/>
  <Default Extension="xml" ContentType="application/xml"/>
  <Override PartName="/word/document.xml" ContentType="application/vnd.openxmlformats-
    officedocument.wordprocessingml.document.main+xml"/>
  <Override PartName="/word/numbering.xml" ContentType="application/vnd.openxmlformats-
    officedocument.wordprocessingml.numbering+xml"/>
  <Override PartName="/word/styles.xml" ContentType="application/vnd.openxmlformats-
    officedocument.wordprocessingml.styles+xml"/>
  <Override PartName="/docProps/app.xml" ContentType="application/vnd.openxmlformats-
    officedocument.extended-properties+xml"/>
  <Override PartName="/word/settings.xml" ContentType="application/vnd.openxmlformats-
    officedocument.wordprocessingml.settings+xml"/>
  <Override PartName="/word/theme/theme1.xml" ContentType="application/vnd.openxmlformats-
    officedocument.theme+xml"/>
  <Override PartName="/word/fontTable.xml" ContentType="application/vnd.openxmlformats-
    officedocument.wordprocessingml.fontTable+xml"/>
  <Override PartName="/word/webSettings.xml" ContentType="application/vnd.openxmlformats-
    officedocument.wordprocessingml.webSettings+xml"/>
  <Override PartName="/docProps/core.xml" ContentType="application/vnd.openxmlformats-
    package.core-properties+xml"/>
</Types>
```

# Openoffice.org

---

- Ingyenes Office alternatíva, az Ms Office után a második legelterjedtebb programcsomag.
- ISO/IEC által szabványosított OpenDocument formátumot használ. A dokumentum kiterjesztése: .ODF
- ODF fájl megnyitása Microsoft Word-ben: csak egy bővítmény segítségével lehetséges (ODF converter)
- Tud DOC-ba, RTF-be és PDF-be is menteni



# Univerzálális formátumok 1.

---

- RTF (Ritch Text Format): A Microsoft fejlesztette ki a különböző verziók közötti kompatibilitási problémák elkerülésére.
- Előnye: szinte minden létező irodai programcsomag képes az olvasására és ilyen formátumba menteni.
- Hátrányai:
  - ha képi elemeket is tartalmaz, akkor nagy méretű lesz a fájl
  - nem képes minden formázási elem megtartására

## Univerzálális formátumok 2.

---

- TXT (Text): a legrégebbi formátum. Formázások csak minimális mértékben valósíthatóak meg (soremelés, tabulátor, stb.)
- Előnyei: kis méret, külön program nélkül olvasható, szerkeszthető
- Hátrányai: Formázások hiánya, nincs egységes karakter kódolás (ékezetek problémája)

# Univerzálális formátumok 3.

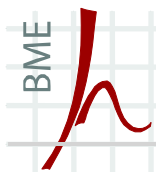
---

- PDF (Portable Document Format): Adobe Systems fejlesztette ki
- A tartalom leírása eszköz és felbontásfüggetlen (mindenhol egységesen fog kinézni)
- Nyílt szabvány
- Előnyei: széles körben elterjedt, minden formázási elemet kezel
- Hátránya: a PDF dokumentumok nem szerkeszthetők (kivéve a PDF űrlap)

# Fájlok konvertálása 1.

---

- Különböző Word verziók, RTF készítése:  
File / Mentés Másként...
- Openoffice.org dokumentum konvertálása  
Word-be:
  - az Openoffice.org alkalmazásban elmentjük  
DOC-ba a dokumentumunkat
  - Bővítmény telepítése a Microsoft Office-hoz:  
<http://odf-converter.sf.net>



# Fájlok konvertálása PDF-be

- Program segítségével (Adobe Acrobat, Openoffice.org, stb.)
- Ingyenes online konvertáló szolgáltatások, pl.:  
<http://www.freepdfconvert.com/>
- PS nyomtatós megoldások
  - „Kézzel”
  - FreePDF (<http://freepdfxp.de>)



# A FreePDF

- Jelenleg a 4-es verziónál tart (ez már nem XP)
- Német fejlesztés, de 16 nyelven beszél
- Feltelepít egy PS kompatibilis nyomtatót
- Redmon-t használ a nyomtatási feladat (PS) fájlba írásához
- GhostScript 8 kell hozzá (ez végzi el tulajdonképpen a PS->PDF konverziót)
- Azaz egy felületet ad és a telepítést/beállítást segíti
- Nem konvertálja a dokumentumokat, „csak nyomtatja”

