

# az univerzális dokumentum létrehozása

PCLinuxOS Magazine – 2022. október

Írta: Paul Arnote (parnote)

Valamikor régen, még az ősi komputerek idején ... esetünkben 30 éve (1992) ... az Adobe létrehozta a PDF fájlt. A kiterjesztés betűi a Portable (hordozható) Dokumentum Fájl"-t takarják. Addig sokkal nehezebb volt dokumentumokat megosztani. Különböző szövegszerkesztő programok voltak forgalomban (*jobbak, mint az MS Word – ford.*) és nem volt garancia arra, hogy az egyikben készített fájl egy másik képes megnyitni. Emellett, ha a megfelelő betűtípus nem volt a fogadó gépen telepítve, a dokumentum nem jelent meg megfelelően.

## Háttér

Szoftveres cégként minden hibája ellenére az Adobe sikerrel hozott létre egy valóban hordozható dokumentum-rendszert, ami minden gépen megfelelően jelent meg. Mindössze egy, a PDF-fájlok kezelésére képes olvasóra volt szükség. Az Adobe az olvasót is elérhetővé tette, és sokáig az volt az egyetlen olvasó a piacon. A PDF-formátum alapját a PostScript-nyelv képezte (amit szintén az Adobe fejlesztett ki 1982 és 1984 között) és minden egyes PDF-fájl egy teljes és kötött szerkezetű sima dokumentumot tartalmazott benne a szöveggel, a betűkkel, vektorgrafikával, raszteres képekkel és más, a megjelenítéshez szükséges információval.



1993-ban az Adobe a PDF fájlspecifikációját ingyen elérhetővé tette. Az első 16 évben a PDF-formátum jogvédett maradt az Adobe kezelésében. És, ahogy emlékszem, az Adobe **keményen** és **vehemensen** felügyelte a PDF-formátumot annyira, hogy elriasztotta a többi szoftveres céget PDF-ek, illetve PDF-olvasók készítésétől. Ezután 2008-ban, az Adobe a PDF-formátumot nyílt szabvánnyá tette, ami egybe esett a Nemzetközi Szabványügyi Szervezet általi nyílt szabvánnyá minősítésével ISO 32000-1 jelzés alatt. 2008-ban az Adobe az ISO 32000-1-re kiadott egy nyilvános szabadalmi licenct, lehetővé téve az Adobe által birtokolt minden, a PDF-

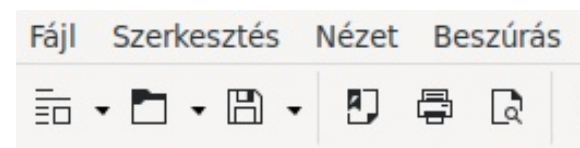
kompatibilis alkalmazások készítéséhez, használatához, eladásához és terjesztéséhez szükséges szabadalom fizetségek nélküli használatát, ahogy azt a [Wikipédia](#) mondja.

Ugorjunk a mába, amikor a PDF dokumentumformátum az iratok meghatározó, de facto szabványává vált. Minden elképzelhető modern platform alatt olvasható és felhasználható. Használatát átvették a számítógépes világ minden sarkában, a hivatalos kormányzati nyomtatványok terjesztésétől kezdve egészen eddig a magazinig. Az irodai csomagok, az MS Office-tól a LibreOffice-on át egészen a Google Docs-ig, lehetővé teszik dokumentum mentését PDF-ként. A PDF-olvasók listája túl hosszú a felsoroláshoz, még a főbb webböngészők is (Firefox, Chrome stb.) képesek PDF-fájlokat natív módon megjeleníteni.

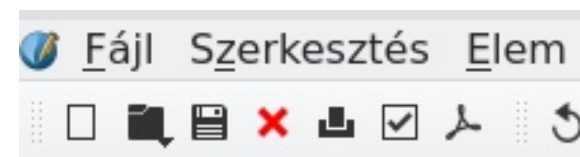
A PCLinuxOS-használók számos opció között válogathatnak PDF-fájl olvasás és megtekintés terén. A KDE-ben ott a kítűnő és ügyes Okular olvasó és néző, nem kizárólag PDF-re, hanem más dokumentumformátumra is. Mate- és Gnome-használóknak ott az Evince, ami csodálatos munkát végez PDF-fájlok pontos megjelenítésében. A Qpdfview egy másik PDF olvasó, néző (nagyon gyakran használom), ami a könnyű súlyával emelkedik ki. Az Evince népszerű Gtk-alapú PDF-olvasó. Sok-sok PDF-olvasó és néző érhető még el, és mindnek megvan a elhivatott tábor.

## PDF-fájl készítése

Amikor a PDF-formátum először megjelent, csak azok tudtak PDF-fájlt készíteni, akik kifizettek az Adobe-nak a PDF-készítő szoftverért egy iszonyatos és szemérmetlenül nagy összeget. Ugyanakkor 2008-tól, amikor a PDF-formátum nyílt szabvánnyá vált, szinte bárki képes lett PDF-dokumentumot készíteni.



LibreOffice Writer ... 4. ikonbalról.

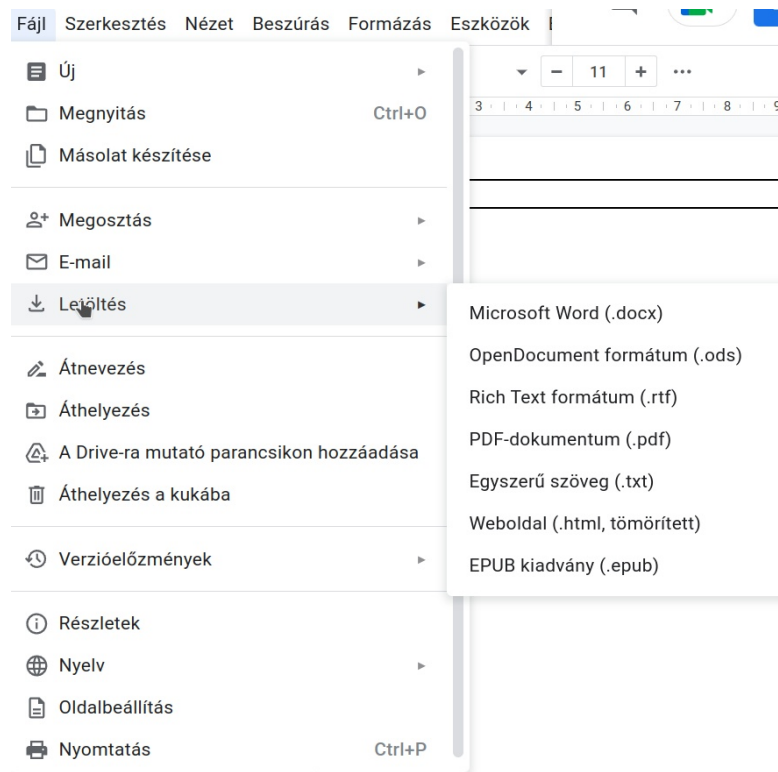


Scribus ... 7. ikon balról.

## PDF első rész: az univerzális dokumentum létrehozása

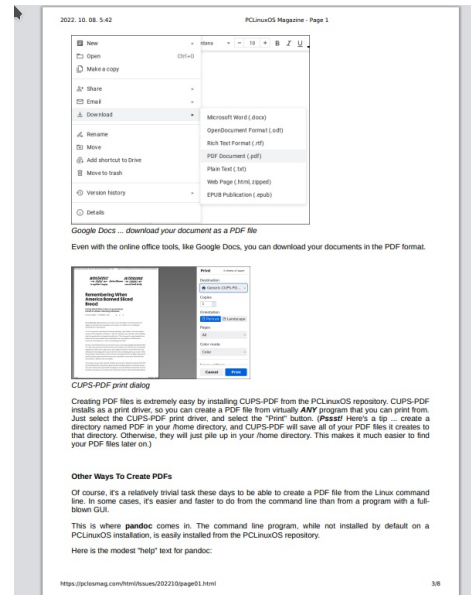
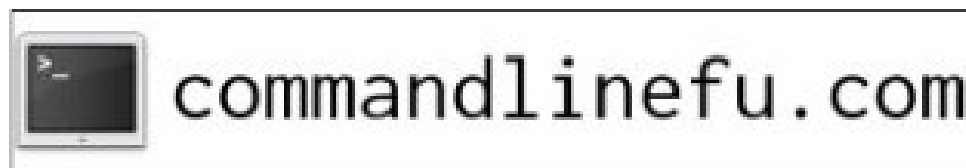
A LibreOffice-nál az eszközsávban (és a fájl menüben) megtalálható a PDF-export ikon, a LibreOffice-csomag VALAMENNYI programjánál. A példában a LibreOffice Writer eszköztárát mutattam. Még a munkahelyemen általam néha (irgalomért könyörögve) használt Microsoft Word és Excel is képes a dokumentumait PDF-fájlként exportálni.

A Scribus (amivel havonta készítjük a The PCLinuxOSMagazine-t – és a fordításokat – a szerkesztő) egy többplatformos asztali kiadványszerkesztő eszköz és az elsődleges kimeneti fájlja a PDF-formátum.



Google Docs ... dokumentum letöltése PDF-fájl formájában

Még az online irodai eszközöknél, mint a Google Docs, a dokumentumok PDF-formátumban is letölthetők



CUPS-PDF nyomtatási párbeszédablak

A CUPS-PDF telepítésével aPCLinuxOS-tárolójából a PDF-készítés kimondottan egyszerűvé válik. A CUPS-PDF nyomtató-meghajtóként települ, így szinte **BÁRMELYIK** programból, ahol nyomtatni lehet, készíthető PDF-fájl. Csak válaszd ki a CUPS-PDF nyomtató-meghajtót és üss a „Nyomtatás” gombra. **(Psztt!Egy tipp! Készíts a /home alatti könyvtáradban egy PDF nevű könyvtárat és a CUPS-PDF az összes PDF-fájlt oda fogja menteni. Ellenkező esetben magában a /home könyvtáradban halmozza fel. Így sokkal könnyebben megtalálhatod a PDF-fájljaidat később.)**

### PDF-készítés másképpen

Természetesen napjainkban szinte magától értetődő, hogy a Linux parancssorában is lehet PDF-fájlt készíteni. Bizonyos esetekben sokkal könnyebb és gyorsabb parancssorból csinálni, mint egy teljes értékű grafikus felületi programból.

Ez az, amikor a **pandoc** színre lép. A parancssori program, noha alaphoz nincs a PCLinuxOS-ben telepítve, könnyen felrakható a PCLinuxOSTárolójából.

Íme a pandoc szerény „súgója”:

pandoc [OPTIONS] [FILES]  
 -f FORMAT, -r FORMAT --from=FORMAT, --read=FORMAT  
 -t FORMAT, -w FORMAT --to=FORMAT, --write=FORMAT  
 -o FILE --output=FILE  
     --data-dir=DIRECTORY  
 -M KEY[:VALUE] --metadata=KEY[:VALUE]  
     --metadata-file=FILE  
 -d FILE --defaults=FILE  
     --file-scope  
 -s --standalone  
     --template=FILE  
 -V KEY[:VALUE] --variable=KEY[:VALUE]  
     --wrap=auto|none|preserve  
     --ascii  
     --toc, --table-of-contents  
     --toc-depth=NUMBER  
 -N --number-sections  
     --number-offset=NUMBERS  
     --top-level-division=section|chapter|part  
     --extract-media=PATH  
     --resource-path=SEARCHPATH  
 -H FILE --include-in-header=FILE  
 -B FILE --include-before-body=FILE  
 -A FILE --include-after-body=FILE  
     --no-highlight  
     --highlight-style=STYLE|FILE  
     --syntax-definition=FILE  
     --dpi=NUMBER  
     --eol=crlf|lf|native  
     --columns=NUMBER  
 -p --preserve-tabs  
     --tab-stop=NUMBER  
     --pdf-engine=PROGRAM  
     --pdf-engine-opt=STRING  
     --reference-doc=FILE  
     --self-contained  
     --request-header=NAME:VALUE  
     --abbreviations=FILE  
     --indented-code-classes=STRING  
     --default-image-extension=extension  
 -F PROGRAM --filter=PROGRAM  
 -L SCRIPTPATH --lua-filter=SCRIPTPATH

--shift-heading-level-by=NUMBER  
 --base-header-level=NUMBER  
 --strip-empty-paragraphs  
 --track-changes=accept | reject | all  
 --strip-comments  
 --reference-links  
 --reference-location=  
     block | section | document  
 --atx-headers  
 --listings  
 --incremental  
 --slide-level=NUMBER  
 --section-divs  
 --html-q-tags  
 --email-obfuscation=  
     none | javascript | references  
 --id-prefix=STRING  
 --title-prefix=STRING  
 --css=URL  
 --epub-subdirectory=DIRNAME  
 --epub-cover-image=FILE  
 --epub-metadata=FILE  
 --epub-embed-font=FILE  
 --epub-chapter-level=NUMBER  
 --ipynb-output=all|none|best  
 --bibliography=FILE  
 --csl=FILE  
 --citation-abbreviations=FILE  
 --natbib  
 --biblatex  
 --mathml  
 --webtex[=URL]  
 --mathjax[=URL]  
 --katex[=URL]  
 --gladtex  
 --trace  
 --dump-args  
 --ignore-args  
 --verbose  
 --quiet  
 --fail-if-warnings  
 --log=FILE  
 --bash-completion  
 --list-input-formats

**-D FORMAT**  
**--list-output-formats**  
**--list-extensions[=FORMAT]**  
**--list-highlight-languages**  
**--list-highlight-styles**  
**--print-default-template=FORMAT**  
**--print-default-data-file=FILE**  
**--print-highlight-style=STYLE|FILE**  
**-v**  
**-h**  
**--version**  
**--help**

Itt a program leírása a Pandoc kézikönyvlapjairól (**man pandoc**)

*A Pandoc egy Haskell könyvtár egyik jelölőnyelvi formátumról másokra fordításhoz és egy parancssori eszköz, ami ezt a könyvtárat használja.*

*A Pandoc képes számos jelölő és szövegfeldolgozó formátum között konvertálni, közte, de nem korlátozva az olyanokra mint a Markdown, a HTML, a LaTeX és a Word docx. A bemeneti és kimeneti formátumok teljes listáját lásd a --from és --to opcióknál lent. A Pandoc PDF-kimenetet is képes előállítani: lásd a PDF-készítést lent.*

*A Pandoc továbbfejlesztett Markdown-ja tartalmaz táblázati szintaxist, definíciós listákat, metaadat blokkokat, lábjegyzeteket, idézeteket, matematikát és sok egyebet. Lásd a Pandoc's Markdown alatt.*

*A Pandoc moduláris felépítésű: válogatott olvasókat tartalmaz, amik elemzik a szöveget meghatározott formátumban és natív megjelenítést készítenek belőle (egy absztrakt szintaxisfát, vagy AST-t), továbbá válogatott írókat, amik ezt a natív megjelenítést a célformátumra konvertálják. Ezáltal, egy új bementi, vagy kimeneti formátumhoz elegendő egy olvasót, vagy író hozzáadni. A felhasználó is futtathat saját pandoc-szűrőket a köztes AST módosítására.*

*Mivel a Pandoc-nál a dokumentum köztes megjelenítése nem annyira összetett, mint amilyen az általa konvertált formátumból több, ne várjunk tökéletes átalakítást az összes formátumról az összes másokra. A Pandoc próbálja megtartani a dokumentum szerkezeti elemeit, de az olyan formázási részleteket nem, mint a margó mérete. Emellett, néhány dokumentumelem, mint az összetett táblázatok, nem biztos, hogy beleférnek a Pandoc egyszerű modelljébe. Miközben törekszik a saját Markdown-járól tökéletes konverzióra, a saját Markdown-nál összetettebb formátumokról konvertálás várhatóan veszteséges lesz.*

Egyszerűsített működésében a Pandoc képes PDF-et készíteni a számítógéped egyszerű (\*.txt kiterjesztésű) szövegfájljából. A Magazin megjelenéséről kiadott értesítésről készítünk egy sima szövegfájlt, benne az összes információval, ami képezi a Magazin kiadási megjegyzéseit.

## PDF első rész: az univerzális dokumentum létrehozása

Így a példa kedvéért, vettem a PCLinuxOS Magazine 2022. júliusi szám megjelenési értesítőjét. Beléptem abba a könyvtárba, ahol azt tároltam és megnyitottam egy terminált abban a könyvtárban.

Ezután kiadtam a következő pandoc parancsot:

**pandoc -o July-2022-Release-Notice.pdf July-2022-Release-Notice.txt**

Ez megmondja a pandoc-nak, hogy a kimeneti fájl (-o parancsopció) azonos lesz a bementivel, de a kiterjesztése .pdf lesz. Ezzel elkészítjük a July-2022-Release-Notice.pdf fájlt a July-2022-Release-Notice.txt bemeneti fájlból. Ami még jobb, hogy több bemeneti fájl is megadható szóközzel elválasztva. Így több bementi fájlból egyetlen kimeneti pdf-fájl is létrehozható, ami magába foglalja az összes bementi fájl tartalmát. Vedd figyelembe, hogy a fájlok felsorolásával a pandoc egyszerűen hozzáfűzi a jobbról második fájlt az első végéhez és a harmadikat a második végéhez és így tovább. A Pandoc alapbeállításként a végső PDF-be egy üres sort illeszt be, mielőtt hozzáadná a következő fájlt az előzőhöz.

The PCLinuxOS Magazine staff is pleased to announce the release of the July 2022 issue. With the exception of a brief period in 2009, The PCLinuxOS Magazine has been published on a monthly basis since September, 2006. The PCLinuxOS Magazine is a product of the PCLinuxOS community, published by volunteers from the community. The magazine is led by Paul Arnote, Chief Editor, and Assistant Editor Meemaw. The PCLinuxOS Magazine is released under the Creative Commons Attribution- NonCommercial-Share-Alike 3.0 Unported license, and some rights are reserved. All articles may be freely reproduced via any and all means following first publication by The PCLinuxOS Magazine, provided that attribution to both The PCLinuxOS Magazine and the original author are maintained, and a link is provided to the originally published article.

In the July 2022 issue:

- Charles H. Tankersley (1933 - 2022): In Remembrance
- PCLinuxOS Recipe Corner: Beef Wellingtons
- GIMP Tutorial: Tool Features You May Not Have Seen
- Short Topix: Gmail Trick Reveals Who Is Selling Your Data
- Facebook Says Apple Is Too Powerful. They're Right
- VLC Channel Scan For USB TV Tuner ... Simplified
- And much more inside!

This month's cover was designed by Meemaw.

*Képernyőkép a Pandoc-kal készített PDF-ről Qpdfview olvasóban.*

A Pandoc nagy tudású kis program. A használatát a lehető legegyszerűbb példán keresztül vázoltam. Ám a Pandoc ennél sokkal több. Számos formátumban képes kimeneti fájlt előállítani. Néhány a leggyakoribb kimeneti formátumból Oo/LO ODT, PDF, HTML, Microsoft Word DOCX, DocBook, DokuWiki, EPUB, FB2, JSON, LaTeX, MediaWiki, és man, hogy csak megnevezek párat. Képes az általa támogatott formátumok között



konvertálni is. A próbáim során nem tudtam kifogni a Pandoc-on konvertálásban, gyorsan és fennakadás nélkül működött. Ám a cikkünk érdekében csak a PDF-fájlokra korlátozzuk magunkat. Ha többet akarsz megtudni a Pandoc más irányú felhasználásáról, akkor javaslom a kézikönyv tanulmányozását az elinduláshoz, olyan módon, hogy a parancssorba beírod a **man pandoc** utasítást.

A Pandoc a kimeneti fájl irányítására, igazítására egy speciális jelölő nyelvet, a Markdown-t használja. A jelölő nyelv használatával meghatározható, hogy a dokumentum melyik része jelenjen meg fékövéren, dőlten szedve és így tovább. Hogy a téma még bonyolultabb legyen, a jelölőnyelvnek számos „formátuma” és „dialektusa” is van. Szerencsére a Pandoc képes sok ilyen jelölőnyelvi dialaktus formátuma közötti konverzióra is. Ha a kíváncsiságod kielégítésére tanulmányozni szeretnéd a témát, indulásképpen [itt van](#) egy leírás a Markdown-ról. De a Pandoc, Markdown-kód nélkül is megcsinálja neked az új PDF-dokumentumodat. Azt azért vedd figyelembe, hogy a Markdown-kód nélkül a kimeneted egy kicsit szegényes lesz. Ám, még így is SOKKAL jobban néz ki, mint az ennél is sokkal fapadosabb kinézetű text-fájl.

Végül, a Pandoc PDF-ek készítésére a LaTeX-et használja. Ezáltal a PDF-ek elég kellemesen néznek ki. A LaTeX használata miatt, a számítógépeden kell legyen valamilyen LaTeX-verzió. Ha a PCLinuxOSTárolójából úgy akarsz telepíteni a Pandoc-ot, hogy nincs a kívánt LaTeX-csomag telepítve (nekem nem volt a Pandoc telepítésekor), akkor függőségként települni fog.

### Összegzés

Gyakran természetesnek vesszük, hogy van. Gondolom azért, mert annyira elterjedté vált a komputeres világában. Ám hol lennének PDF-fájlok nélkül?

A PDF fájlformátum RENGETEG jó dolgot tartalmaz. Az első természetesen az univerzális hordozhatóság a

különbéféle számítógépes platformok között (igen, még a DOS-nak is van PDF-olvasója).

A második, hogy a PDF-fájlok tokenezett PostScript-nyelvet használnak a dokumentum megfelelő formátumának megtartásához, miközben még a betűtípusokat is tartalmazzák. Ez biztosítja, hogy függetlenül a PDF megtekintésekor használt platformtól egyformán fog megjelenni.

A harmadik, hogy a PDF-fájlok egy kicsit biztonságosabbak, mintha simán küldenénk egy dokumentumfájlt, amit nagyon könnyű módosítani. Lehet a PDF-eket szerkeszteni? Igen, lehet. Ám ez nem annyira egyszerű, sokakategyáltalán nem érdekel.

Az elkövetkező cikkekben mélyebbre hatolunk a PDF-ek manipulálásának és szerkesztésének világába. Ezen feladatok egy része parancssori eszközökre támaszkodik, mások grafikusra. Ezek során megismerkedünk pár új linuxos tippel, trükkel és eszközzel.

