# A digitalizált mikrofilmek szolgáltatásának egyik alternatívája: a kétrétegű PDF

2010. február 4-én az Országos Széchényi Könyvtárban előadást tarthattam a helyi lapok digitalizálásának lehetőségeiről, a kétrétegű PDF-ek készítésének mikéntjéről, illetve ezek könyvjelzőzésének lehetőségeiről. Bár az eljárásban nincs semmi új és meglepő, talán érdemes írásban is összefoglalni az ott elhangzottakat, hogy 1. az itt leírtak kellően felboszszantsák az e kérdésekhez jobban értőket és hozzászólásaik nyomán kollégáim újabb ismereteket szerezhessenek, és/vagy 2. választ adjanak az előadáson el nem hangzott, vagy az azóta hozzám érkezett kérdésekre.

# Miről is van szó?

Néhány városi könyvtárnak – egy állományvédelmi digitalizálási program keretében - módja nyílt az Országos Széchényi Könyvtártól (OSZK) megrendelni a helyi lapjaikról készített mikrofilmek digitális másolatait, illetve megvásárolhatta az ABBYY FineReader (Professional Edition) optikai szövegfelismerő rendszer 9.0-ás, valamint az Adobe Acrobat Pro szintén 9.0-ás verzióját. A támogatás elnyerésének feltétele volt, hogy a számítógépes programok segítségével a hírlapokról, folyóiratokról készült képeket használatra, kutatásra tegyék alkalmassá, tegyék lehetővé munkájuk eredményéhez a minél szélesebb körű hozzáférést.

Mi tehát a teendő a programok megvásárlása és a mikrofilmek digitális másolatainak megrendelése, megérkezése után?

### Nézzük meg, mink van (és hogy veheti hasznát szinte azonnal az olvasó)

Ahhoz, hogy lássuk, mit kaptunk az OSZK-tól, először is szükségünk lesz egy képnézegető programra. Erre a célra az ingyenesen, magyar nyelven is használható IrfanView programot ajánlom<sup>1</sup>.

Ha e programmal megnyitjuk és megnézzük a képeket<sup>2</sup>, először is azt tapasztaljuk, hogy egy kép két újságoldalt tartalmaz és ennek következtében szükségszerűen találunk olyan képeket, amelyek két lapszám egy-egy oldalát tartalmazzák. A képek tulajdonságairól a képnézegető program az i betű lenyomása után tájékoztat (1. ábra).

Az így nyert információkból megállapíthatjuk, hogy 300 DPI felbontású, 4784x3560 pixel méretű, kétszínű, TIF-képek állnak rendelkezésünkre.

Mit tegyünk, hogy azonnal szolgáltatni tudjuk szerzeményünket? Először is a megvásárolt képekről készítsünk mentést és gondoskodjunk az elmentett állományok biztonságos tárolásáról. Ezután rendezzük el a fájlokat egy, az olvasók számára is áttekinthető struktúrába, például így: Létrehozunk egy alkönyvtárat a lap nevével, és ebbe évenként, folyamatosan számozva helyezzük el a képeket:

nagy_kunsag <sup>3</sup>	
1885	
	1885_001.tif
	1885_002.tif
	1885_003.tif
1886	
	1886_001.tif
	1886_002.tif
	1886_003.tif

. . .

(Ehhez célszerű egy "commander" programot telepíteni és megtanulni a fájlok csoportos átnevezésének módját.) A sorszámozásnál mindig figyeljünk a kellő számú "vezető nulla" használatára (001, 002, 099, 999), hogy fájljaink sorba rendezése megfelelő legyen!

Olvasóink így – a mikrofilmmel szemben – egy kényelmesebben használható, nyomtatott és elektronikus másolatot is nyújtani tudó szolgáltatáshoz jutnak, ami - lássuk be - nagy előrelépés!



1. ábra Képinformációk megtekintése az IrfanView programban

E szolgáltatásunk persze még sok kényelmetlenséget okoz. Az olvashatóság érdelében a képernyőn jelentősen fel kell nagyítani az oldalak képét, ez zavarja az áttekinthetőséget és ebből a szempontból (sem) szerencsés, hogy egy kép két újságoldalt is tartalmaz (2. ábra).

🍁 1885_006.bf - IrfanView (Zoom: 2816 x 2096)	
Eajl Szerkesztés Kép Beállítások Nézet Súgó	
😹 🖬 🗙 🕹 🗐 🛈 💐 🍕 🗢 의 Ol Ol 645 📓 🕐	
and the second	A CONTRACT OF
a south and a second as a second s	and the second se
irám, az Hiszen a vidéki vagy az idegen a ki a	and the second
kiállitásra jön, a városligetben a kimerülésig	akár társaságot formálva, akár más községgel   szo
k itt sey- elfarad a sok latni-valoban, örul tehat, hogy	ctatlakozva, vagy önállóan megvéve e birtokot kis
arkedas tat a nagyhöségben nem kénytelen a városba visz-	használják birtokosaiuk, ez altal csak jót fognak
ie, és a ta- sza jöuni, hanem minden igényeit, minden va-	teoni, a mi nekik is bizonyosan használni fog, s Sz
melyike sen. gyat a hüs szabadban, szemet kapraztato latm	ha az öseinkról reank maradt vagyonokat fenttar- tiz
es r 100 valor rozepette. Egy kisse dragauoun, az igaz,	tani, vagy gyarapitani fogjuk ez által a jövőnem- va
, de a kun- de a kenyelem es ralodi elvezet miatt mar	zedek előtt is érdemünket tartjuk fel, mig e va- a
laitam ada- nem fukarkodik,	gyonok eitdegenitésével és sajit czéljainkra fel- ba
načamat e A klalitasi vendeglosoknek nem is ienet	hasznalásával, csak a későbbi kor kárhoztatását tá:
egynazior- paaasaa, mert a nodmezovasarnetyi csaruat	vonnok magunkra, s ez alapon az elidegenitésre, ge
m nyernet- siveve, minden nety napongent zsatotasig tonive	az erkolcsi zotelesseg szemponijaból is, a midőn ké
hie harr - a keleti ér harmált kárábia zamatar tá	der kein tartanunk azt a mivel rendelkezünk tulaj- eli
staint boy, rok béndival de ingerlő áltórató ningedenőival	doukeppen jogunk nincs. Azert cak udvozolhetjuk 18
aformatió a szumunal de excinale la talmas sousará. Tábban	a bemes eszinecserer, s elismeresünk illeti a birt. gy
is ezt låt. zarändekolnak oda mint sek kielliteti pa-	berdenel machantlation blittel at minden ily gy
által levén i villonha	nest des a bistokenda alignett alignet alignett
nézis csak Mert hizony yan yany 8-10 olyan na.	adentical terbesting star is minin bankers a
n történeti gyobb pavillon a melyik kong az ürességtől a	kardaus hirtok a karal inghan min anna dalah (a
a melynek látogatói tisztán csak szakemberek.	laddk altal is marter neverie ait alranot is man
sigunk re- Annál keresettebb a nagy iparcsarnok, a ke-	Vannak mar av Barnak avarna kik alabh ant a
negirni, ha leti és bosnyák pavillonok, mely két utóbbi-	karida hirtalan falvatáránál - nam tastatták salas
mely egy ban jó vásárt csinálnak a kiállítók, lévén ott	elónyösnek. Mi bizunk egen Beyek id és fidete bi
ullast, kik minden tárgy eladó. A kjállitási emlékek és	menetelében. A birt, gyülés tárgyát képezte még
k lehettek ajándékok is jobbadán innen kerülnek ki, mi-	s Paka pusztai megvizszálatlan számadánnak a mit
iskolák. E ntán ezek a keleri tárovak naovan dedaleseale	an hatestidatet tones an ide manificat faith fat
·	· ·
4764 x 5360 x1 8FP 0/03 59 % 738.84 K8 / 2.04 M8 2009.07.02. / 07:40:56	

2. ábra Az eredeti képek áttekintésének nehézségei

Az információkhoz való hozzáférést az segíti igazán, ha a képeinken látható szöveget kereshetővé tesszük. Az első problémát (oldalak szétvágása) ugyan megoldhatnánk képnézegető programunkkal, egyéb célszoftverrel is, de ne tegyük. A megvásárolt karakterfelismerő program 9.0-ás verziója ezt a problémát is meg fogja oldani.<sup>4</sup>

Indítsuk el és győződjünk meg néhány beállításáról!

#### Kétrétegű PDF készítés

Az ABBYY FineReader 9.0 használata során rengeteg beállítási lehetőségünk van. Ezek közül számunkra most kettő kiemelten fontos. Ahhoz, hogy a két oldalt tartalmazó képek szétvágásával ne kelljen foglalkoznunk, az "Eszközök" legördülő menü "Beállítások" pontjára kattintás után megnyíló ablak "1. Beolvasás/Megnyitás" fülén (az egérrel) tegyünk egy pipát "A dupla oldalak felosztása" felirat előtti négyzetbe (3. ábra).

#### TMT 58. évf. 2011. 1. sz.



3. ábra Az ABBYY FineReader 9.0 beállítása (1)

Ahhoz, hogy kétrétegű PDF-ünk (a két réteg: a kép és a szöveg) első, látható rétege a dokumentum *képe* legyen, a felismert szöveg pedig a háttérben maradjon, a "3. Mentés" fül "PDF" alfülén szintén győződjünk meg néhány beállításról (*4. ábra*).

	2. Erte	elmezés	3. Mentes	Nézet	Speciális	1	
RTF/DOC/WordML/DOCX XLS/XLSX	PDF	HTML	PPT/PPTX	TXT	DBF	CSV	LIT
Alapértelmezett papírméret		774			- 00 		-
Az eredeti képméret megtart	•						
Mentési mód	_						_
Szöveg az oldalkép alatt	-						
A bizonytalan szavak helyettes	ítése ké	pekkel					
A saines de béttérenie meste							
A SZOVEG- ES Hatterszin megta	tasa						
A feilécek és a lábienyzetek megtal	rtása entartás	a					
A fejlécek és a lábjegyzetek megtal     Címkézett PDE engedélyezése	rtása egtartás (Acroba	a t50 var	ny újabbal koj	matihilis	-)		
A szöveg- es nattergan megtal     A fejlécek és a lábjegyzetek me     Gímkézett PDF engedélyezése	rtasa egtartás (Acroba	a t 5.0 vag	gy újabbal ko	mpatibilis	5)		
<ul> <li>A szöveg- és nattergan megan</li> <li>A fejlécek és a lábjegyzetek me</li> <li>Címkézett PDF engedélyezése</li> <li>Vegyes rasztertartalom haszná</li> </ul>	rtasa egtartás (Acroba ilata	a t 5.0 vaç	gy újabbal ko	mpatibilis	5)		
A szöveg- es hattejzin negal A fejlécek és a lábjegyzetek m Qímkézett PDF engedélyezése Vegyes rasztertartalom haszná Képbeállítások	rtása egtartás (Acroba ilata	a t 5.0 vaç	gy újabbal ko	mpatibilis	5)		
A szövég-es naue gan neguai A fejlécek és a lábjegyzetek mi Gjmkézett PDF engedélyezése Vegyes (asztertartalom haszná Képbeállítások Egyéni	rtasa egtartás (Acroba ilata	a t 5.0 vaç	gy újabbal ko	mpatibilis	5)		
A szövég-és nakté gan negyai A fejlécek és a lábjegyzetek mi Gjmkézett PDF engedélyezése Vegyes (asztertartalom haszná Képbeállítások Egyéni Betűtípusok beállításai	rtasa egtartás (Acroba ilata	ia t 5.0 vaç	gy újabbal ko	mpatibilis	s)		
A Szöveg-'es nakde gan meguan A fejickek és a lábjegyzetek me Gímkézett PDF engedélyezése Vegyes (asztertartalom haszné Képbeállítások Egyéni Betűtűpusok heallításai Alapbetűtípusok használata	rtasa egtartás (Acroba ilata	a t 5.0 vag	gy újabbal ko betűtípusok bi	mpatibilis ságyazá	s) sa		
A Szövég-es nauteigan meguai A égiécké sa lábigeyzetek me Gímkézett PDF engedélyezése Vegyes (asztertartalom haszné Képbeállítások Egyéni Betűtípusok beállításai Alapbetűtípusok használata Biztonság	rtása egtartás (Acroba ilata	a t 5.0 vaç	gy újabbal ko vetűtípusok bi	m <b>patibili</b>	s) sa		
A Szövég-'es naude gan megua an A Egiécké sa lábigeyzetek me Címkézett PDF engedélyezése Vegyes (asztertartalom haszná Képbeálítások Egyéni Betűtípusok beállításai Alapbetűtípusok használata Biztonság PDF hytonséni beállításai	rtasa egtartás (Acroba ilata	ia t 5.0 vag	gy újabbal ko vetűtípusok bo	mpatibilis eágyazá	s) sa		
A 520VEY'ES nauce gan nergua Q A Egickek és a lábjegyzetek ne Q Cinkézett PDF engedélyezése Vegyes rasztertartalom haszná Képbeállítások Egyéni Betűtípusok beállításai Alapbetűtípusok használata Biztonság PDF biztonsági beállításai	rtása egtartás (Acroba ilata	ia t 5.0 vag A b ság nélkü	gy újabbal ko vetűtípusok bo	mpatibilis eágyazá	s) sa		

4. ábra Az ABBYY FineReader 9.0 beállítása (2)

Itt figyeljünk arra, hogy az "Alapértelmezett papírméret" pontnál "Az eredeti képméret megtartása", a "Mentési mód"-nál pedig a "Szöveg az oldalkép alatt" opció legyen kiválasztva.

Ha e két beállításról gondoskodunk, akkor képeink betöltése után<sup>5</sup> a következőket látjuk (*5. ábra*).

A program tehát – beállításunk nyomán – automatikusan elvégzi a képek szétvágását. Ezeket a képeket a "Fájl" – "Képek mentése" menüpont segítségével feltétlenül mentsük el és nevezzük át őket, hogy olvasóink – addig is, míg munkánk végére nem érünk – a már említettnél kényelmesebben használható szolgáltatáshoz jussanak. (6. ábra).



5. ábra Az ABBYY FineReader használata

<u>F</u> ájl	Szerkesztés <u>N</u> ézet <u>D</u> okumentum <u>O</u> ldal	B <u>l</u> okkok	Eszk
	Oldalak beolvasása	Ctrl+k	(
	PDF fájl/kép megnyitása	Ctrl+C	)
	FineReader dokumentum megnyitása	Ctrl+Shift+N	1
	Új FineReader dokumentum	Ctrl+N	1
	FineReader dokumentum mentése		
	Mentés másként		•
	Küldés		•
	Kijelölt oldalak küldése		٠
	Képek mentése	Ctrl+Alt+S	5
	Postázás		•
	Nyomtatás		•
	FineReader dokumentum bezárása		
	Legutóbbi dokumentumok		
	Kiilépés		

6. ábra Képek mentése az ABBYY FineReader 9.0 program segítségével

A képek jobb használhatósága érdekében a következő fájlszerkezet kialakítását javaslom: nagy\_kunsag

...

1885 1885-05-31\_01.tif 1885-05-31\_02.tif 1885-05-31\_03.tif 1885-05-31\_04.tif

1885-06-07 01.tif

Azaz: megjelenés éve-hónapja-napja\_sorszám. kiterjesztés. Itt se feledkezzünk meg a "vezető nullákról". Fontos: a képeket úgy nevezzük el, hogy a majdan használandó, keresésre alkalmas PDF állományokban beazonosíthatók és igény esetén visszakereshetők legyenek, mert e képekből jobb minőségű nyomtatásokat lehet készíteni, mint munkánk leendő végeredményéből.

A dokumentumok szövegének felismertetéséhez kattintsunk a "Dokumentum értelmezése" gombra és várjunk (ill. foglalkozzunk egyéb feladatainkkal), míg a program végez a betöltött képek felismerésével...

Ez messze nem lesz hibamentes. A képernyő bal oldalán található kis képek bal alsó sarkában megjelenő, a felismerés befejezését jelző ikon mellett, a jobb oldalon feltehetően sűrűn látni fogjuk a felismerés hibáira utaló felkiáltójelet, a képernyő közepén pedig feltűnik a hibalista is (*7. ábra*).

Mivel a képek minőségének javítására nincs módunk, ebbe törődjünk bele.<sup>6</sup>



# 7. ábra Hibaüzenetek a karakterfelismerő programban

A felismertetés után a FineReader programban lehetőségünk nyílik a szöveg korrektúrázására. Ennek felvállalását azonban alaposan meg kell fontolni, mert irdatlanul nagy munkát vállalnánk vele. Végiggondolhatjuk, kipróbálhatjuk a szöveg meghatározott részeinek (címek, alcímek, szerzők neve, szövegben szereplő nevek stb.) korrektúrázásának lehetőségeit is, mert ez később megkönynyíti a könyvjelzőzés folyamatát.<sup>7</sup> A szövegfelismerő programok feltételezhetően még sokat fognak fejlődni, remélhetőleg előbb-utóbb alkalmasak lesznek 19. századi szövegek helyesebb felismerésére is, illetve talán előbb-utóbb lehetőség nyílik arra is, hogy hírlapjaink mikrofilm-másolatai helyett az eredeti kiadványokról szerezzünk be - a másolat másolatáról készítettnél – sokkal jobb minőségű képeket.

Abban az esetben, ha egy lapnak csak egy-két évfolyama jelent meg egy településen, fel lehet vállalni ezt a sziszifuszi, de kétségtelenül nagyon érdekes feladatot is.

A felismert oldalak kétrétegű PDF fájlokba mentése nagy odafigyelést igényel. Érdemes lapszámonként készíteni a mentést oly módon, hogy a Ctrl gomb lenyomása mellett egérkattintással kijelöljük az egy fájlba menteni kívánt oldalakat. (Ha helyesen végeztük el a fájlok elnevezését, ezek egymás alatt lesznek!) A "Mentés" gomb megfelelő beállításával a következő ábra szerint lehet elvégezni ezt a feladatsort (*8. ábra*).



8. ábra Az ABBYY FineReader 9.0 beállításai a PDF dokumentum mentésekor

#### TMT 58. évf. 2011. 1. sz.

Amire figyelni kell: helyes legyen a fájlnév (ez lehet a megjelenés éve-hónapja-napja), a "Kijelölt oldalak" előtti rádiógomb legyen bekapcsolva, és a "Fájltulajdonságok"-nál az "Egyetlen fájl létrehozása az összes oldalhoz" opció legyen beállítva.

A már elmentett oldalaknál a felismerést jelző ikon mellett megjelenik a mentést jelző ikon is (9. ábra).



9. ábra Az értelmezést és mentést jelző ikonok

Nincs más dolgunk, mint az egyenként elmentett lapszámokat egy alkönyvtárba összegyűjteni és nem megfeledkezni arról, hogy a PDF fájlok használatára ingyenesen elérhető Adobe Reader program "Bővített keresés" funkciójával (Shift+Ctrl+F) az egy alkönyvtárban összegyűjtött összes, meg nem nyitott fájlban is lehet keresést végezni. A bővített keresés a későbbiekben létrehozandó könyvjelzőkre is kiterjed. Ne felejtsük el erre felhívni olvasóink figyelmét! (*10. ábra*).

## A könyvjelzőzés

Már jeleztem, hogy a szöveg felismerése nem 100%-os. A könyvjelzők készítésének célja épp ezért az, hogy a szövegben az általunk kijelölt elemek szó szerint és betűhíven kereshetők legyenek a dokumentumban, illetve hogy ezekhez a pontokhoz gyorsan "oda tudjon ugrani" a majdani felhasználó.

Nagy körültekintést kíván a könyvjelzőzendő elemek kiválasztása. Az alábbi példában a címek, a szerzők neve és a rovatcímek kerülnek be a könyvjelzők közé, de az elemek köre természetesen igény szerint bővíthető.

Maga a munka végtelenül egyszerű. Kétrétegű PDF dokumentumainkat megnyitjuk az Adobe Acrobat programmal, a megfelelő ikonra kattintva láthatóvá tesszük a könyvjelző eszköztárat, a képernyőn kijelöljük azt a szövegrészt, ahová könyvjelzőt szeretnénk elhelyezni, majd megnyomjuk a Ctrl+B billentyűkombinációt... (*11. ábra*).

Keresendő:	
papp C:\Users\Takats	
Bela Documents (1885 nelyen	
Eredmények:	
5 dokumentumban 60 elofordulas	
Új keresés	
Eredmények:	
🗆 😼 C:\U\1885-05-31.pdf	*
🛛 🕸 Papp Sándor körútja. Nem i	
🚽 vérnek <b>Papp</b> Sándornak kör	
🕸 mely Papp Sándor zászlóit l	
🕸 n Papp Sándor jelöltet, felsz	
🐳 Ezután Papp Sándor kezdeti	ш
🐟 ajánlotta Papp Sándort. Bég	
widőn Papp Sándor karezag	
😼 Papp Sándor a kun nép igaz	
😼 szavakkal Papp Sándort, a k	
Papp Sándort, s ajánlva őt a	
Papp Sándor után Pap Elek	
widőn Papp Sándor jelölt a	
wutátiin Papp Elek, hogy a Uu	
bogy Papp Sándor a kunok	+
4 III +	
	-

10. ábra Bővített keresés az Adobe Reader programmal



11. ábra Könyvjelzők készítése Adobe Acrobat programmal

A képen látszik, hogy például a "T Á R C Z A ." könyvjelzőt kell átszerkesztenünk TÁRCZA formára, mert a keresés csak így fogja megtalálni. Ha az egeret az elkészült könyvjelző fölé visszük és a jobb oldali gomb megnyomása után a lokális menüből a "Tulajdonságok" pontot választjuk, a különböző kinézetű könyvjelzők megkülönböztetésére is lehetőségünk lesz (*12. ábra*).



#### 12. ábra A könyvjelzők megjelenítési lehetőségei

Ez esetben a szerző nevét félkövér, a rovatcímet dőlt, a címet normál betűvel különböztettem meg.

Ha végeztünk az egyes lapszámok könyvjelzőzésével, a PDF fájlok összevonásával a könyvjelzők listája is összevonható.

Ezt a feladatot szintén az Adobe Acrobat programmal lehet elvégezni. A "Fájl" – "Egyesítés" – "Fájlok egyesítése egyetlen PDF fájlban" lehetőség kiválasztása után a megjelenő ablakban (bal felső sarok) a "Fájlok hozzáadása…" pontra kattintás után kijelölhetjük azokat a lapszámokat, amelyeket egyetlen állománnyá szeretnénk összevonni. Az összevont fájl az Adobe Reader programban így jelenik meg (*13. ábra*).



13. ábra Az összevont könyvjelzők megjelenése

Látszik, hogy az egyes lapszámokhoz tartozó könyvjelzőket (a példadokumentumban nem minden könyvjelző készült el) a fájlok neve fogja egybe, a névadásra tehát tényleg érdemes odafigyelni és a választott formát következetesen alkalmazni! Az így létrehozott fájl mérete természetesen már jóval nagyobb lesz, összevonáskor tehát (az átlagos terjedelem és számítógépeink minőségének függvényében) laptípusonként érdemes meghatározni az összevonandó lapszámok mennyiségét (havi, negyedéves, féléves, éves összesítés).

Az így összevont fájlokat – méretük miatt – interneten keresztül már nem érdemes szolgáltatni.

#### Még néhány szó...

A kétrétegű PDF dokumentumok lokális kereséséről (Adobe Reader, bővített keresés) már esett szó. Ha a szolgáltatás az internetre kerül – ingyenes megoldásként – a kereséshez szeretném mindenki figyelmébe ajánlani a Google "Egyéni kereső" szolgáltatását<sup>8</sup>, amivel – ha nem is 100%-os, de – korrekt szolgáltatás készíthető.

A hálózaton közreadott PDF fájlokban célszerű rögzíteni a dokumentum tulajdonságait. Ezt az Adobe Acrobat program "Fájl" menüjének "Tulajdonságok..." pontjával (vagy a Ctrl+D billentyűkombináció lenyomása után) az ábrán látható ablakban tehetjük meg (*14. ábra*).

írás	Biztonsá	g Betűtípusok	Kezdeti nézet	Egyéni	Speciális		
Leírás							
	Fájl:	1885-05-31_m					
<u>C</u> ím:		Nagy-Kunság. Társadalmi, ismeretterjesztő- és szépirodalmi közlöny					
	Szerző:	Királyvölgyi Vár	si Könyvtár				
	<u>T</u> árgy:						
K <u>u</u> lc	sszavak:	Karcag 1885. május 31. 1885. június 7. (1 1885. június 14. ]	(10. évf. 23. szár 10. évf. 24. szám (10. évf. 25. szár	n) ) n)			
Léti	rehozva:	2010.11.01. 13:20	:42			Kiegészítő metaadat	
Módosítva		2010.11.10. 20:39:53					
Alka	almazás:						
Specia	ális						
PDF	létrehozój	ja: ABBYY FineR	eader 9.0 Profe	ssional E	dition		
	PDF-verzi	ó: 1.4 (Acrobat	5.x)				
	He	ly: F:\files\tmt\j	pdf\1885\				
	Fájlmére	et: 1,57 MB (164	15 754 bájt)				
	Oldalmére	et: 8,05 x 11,93 H	ivk		Oldalak szám	ia: 4	
Cín	nkézett PD	IF: Igen			Gyors webnéze	et: Nem	

14. ábra A dokumentum tulajdonságainak rögzítése

Az itt rögzíthető adatok köre kiegészülhet a "Kiegészítő metaadat..." gombra kattintás után (*15. ábra*).



15. ábra Kiegészítő metaadatok

A metaadatok rögzítése lehetővé teszi, hogy az internetes keresők a találati listában megfelelően jelenítsék meg dokumentumaink adatait, illetve, hogy a felhasználók a letöltött fájlok forrását, lelőhelyét később is azonosítani tudják. Fontos tudni, hogy ha az egyes lapszámok PDF állományait egyenként látjuk el ezekkel az adatokkal, a fájlok összevonása után a sorrendben legelső szám adatai fognak az új, összevont dokumentum egészére vonatkozni, az adatokat tehát ilyen esetben újra kell szerkeszteni.

A metaadatok rögzítésére a fentiekben természetesen csak egy példát mutattam be. Célszerű erre helyi szabályzatot alkotni és a szerint – de mindenképpen következetesen – eljárni addig, amíg ezek kezelésére könyvtárainkban ki nem alakul egy együttes gyakorlat, vagy nem születik erre egy megfelelő ajánlás.

Biztos vagyok abban, hogy a régen megjelent helyi lapok internetes közzététele nagy érdeklődést fog kiváltani a városi könyvtárak honlapjain. Tudom azt is, hogy egy ilyen digitalizálási műveletsort sokkal könnyebb megmutatni, mint leírni, de kedvcsinálónak talán megfelelnek a leírtak... Internetes fórumainkon, személyes kapcsolatainkon keresztül, a jó példák tanulmányozásának segítségével sokat tanulhatunk egymástól. Használjuk ki ezeket a lehetőségeket!<sup>9</sup>

### Jegyzetek

- <sup>1</sup> Letölthető pl. innen: <u>http://www.irfanview.com/</u>. A cikk írásakor e program 4.27-es verzióját használtam.
- <sup>2</sup> Előadásomhoz a karcagi Városi Csokonai Könyvtártól kaptam közel száz képet a Nagy-Kunság c. lap 1885ben megjelent számaiból. Írásom elkészítésekor is ezeket használtam, itt is megköszönve a karcagi kollégák szíves segítségét.
- <sup>3</sup> A könyvtáraknak, fájloknak természetesen adhatunk ékezetes neveket is. A későbbi – interneten is elérhető – szolgáltatás zavartalansága érdekében azonban jobb, ha hozzászokunk az ékezet nélküli megnevezésekhez.
- <sup>4</sup> A képek szétválasztásának természetesen egyéb módjai is vannak. A feladat megoldható pl. a http://scantailor.sourceforge.net helyről letölthető, ingyenesen használható Scan Tailor programmal is. Valamilyen alternatív megoldásra valószínűleg szükségünk lesz, mert nem biztos, hogy a FineReader minden képünket megfelelően fogja kettévágni. Ha csak egy-két esetről van szó, a szétválasztás a képnézegető programmal is megoldható, de ez lassabb, körülményesebb megoldás.
- <sup>5</sup> A betölthető képek mennyisége számítógépünk kapacitásától függ. Egy alkönyvtár összes képét a Ctrl/A billentyűkombinációval tudjuk kijelölni a megnyitáshoz.
- <sup>6</sup> Elvileg lehetne próbálkozni a kontraszt és a fényerő változtatásával az IrfanView csoportos konvertálás funkciójával, de ezeknél a gyenge mikrofilmes képeknél nemigen segít. Viszont a felismerési nyelvre vonatkozó figyelmeztetésekre érdemes odafigyelni, mert lehet, hogy idegen nyelvű (pl. német) rész van valamelyik újságoldalon, és ha mi csak a magyar nyelvet állítottuk be, akkor a FR nyilván hibásan ismeri fel ezt a részt.
- <sup>7</sup> E cikk szerzője sosem a FineReader program segítségével korrektúrázott, másoknak talán ez jobban "kézre áll".
- <sup>8</sup> L. <u>http://www.google.hu/cse/</u>
- <sup>9</sup> E cikk szerzője is sokat tanult Drótos László lektori megjegyzéseiből, javaslataiból, melyek legtöbbjét be is építette írásába. Köszönet érte!

Beérkezett: 2010. XI. 21-én.



# Takáts Béla

a Jász-Nagykun-Szolnok Megyei Verseghy Ferenc Könyvtár és Művelődési Intézet szakmai igazgatóhelyettese, az MKE Helyismereti Könyvtárosok Szervezetének elnöke. E-mail: *tbela*@vfmk.hu