



Molnár Sándor

Folyóirat-digitalizálási projektek a Széphalomtól a DélmagyarArchívig

Bevezetés

Az SZTE Klebelsberg Könyvtár már több mint 15 éve foglalkozik digitalizált folyóiratfüzetek szolgáltatásával, ennek ellenére mindeddig nem született olyan írás, amely összefoglalja a könyvtár ez irányú tapasztalatait. Kevés primer forrás áll a rendelkezésemre, javarészt szóbeli elbeszélések, cikk- és tanulmányrészletek, illetve Networkshop-előadások alapján lehet rekonstruálni a korábbi évek eredményeit. Írásomba a jelenleg futó legnagyobb digitalizáló projektet, a *Délmagyarország* napilap digitalizálását, és a szolgáltatás kialakítását szeretném bemutatni.

A kezdetek

Az Egyetemi Könyvtárban már a 2000-es évek elején voltak olyan kísérletek, amelyek teljes szöveggel elérhető könyvek online szolgáltatását célozták. Az első projekt alanya, a *Széphalom* című szegedi folyóirat volt. A *Széphalom* 1927–1944-ig működött, egyetemtörténeti jelentősége abban áll, hogy a két világháború közötti időszakban ez volt a legjelentősebb szegedi tudományos igényű lap. A szerzők között egyaránt megtalálhatók az egyetemhez (pl. *Szekfű Gyula* és *Trencsényi-Waldapfel Imre*), és a város irodalmi életéhez kötődő személyek (pl. *Juhász Gyula* és *Radnóti Miklós*) is. „A magyar tudományos és irodalmi élet számos kitűnősége körükből származik”¹, ezáltal az országon belül tudományos viszonylatban ez az egyik legjelentősebb szegedi folyóirat.

A megvalósult szolgáltatás, a 2000-es Networkshop konferencián került bemutatásra², a „digitális

Széphalom” a könyvtár honlapjának egyik oldalán kapott helyet.³

A Széphalom átadása után egy évvel készült el a folyóirat teljes szövegű adatbázisa. A szöveg indexelése, egy a könyvtártól független szereplő, az IndexData által fejlesztett Zebra szoftverrel történt⁴. A digitalizált Széphalom korát megelőző szolgáltatás volt, a weben keresztül elérhető volt a folyóirat teljes szövege. Sajnos ma már nem érhető el az adatbázis, mert a könyvtár 2004-es költözésekor megsérült az a szerveregység, ahol ezt az adatbázist tárolták, és a későbbiekben már nem építették újra.

A 2000-es évek elején indult el a Bodza, az Egyetemi Könyvtár „házon belül” fejlesztett rendszere. Indulásának története szorosan összekapcsolódik, a *Bibliotheca Eruditionis* projekttel, amely munka 2003-ban indult⁵, majd 2008-ra készült el a MOKKA-R⁶.

Ezzel párhuzamosan a szegedi Egyetemi Könyvtárban kísérletek indultak különböző Bodza-alapú folyóirat-indexelési projektek megvalósítására, a *Földrajzi Közlemények* (2002), a *Harmadkor*⁷ (2008), a „Dolgozatok a Magyar Királyi Ferencz József Tudományegyetem Archaeologiai Intézetéből” (2009). E három kiadvány sok szempontból rokonságot mutat egymással, teljes szövegű indexelés egyik esetében sincs, de lehetőség van a digitalizált állomány képként való elérésére. Ebben a három adatbázisban nincsenek letölthető PDF fájlok, illetve nincs OCR-ezett szöveg sem. A tartalom visszakereshetőségét a hagyományos könyvtári címleírás adja.

Az 1. ábrában látható, milyen „lépcsők” után jutott el a könyvtár a jelenleg használt szolgáltatási helyek kialakításához. 2010 előtt nem volt evidencia

digitalizált anyagaink fulltext indexelése, noha az első kísérlet már 2002 környékén megtörtént.

A Bodza alapú adatbázisok tapasztalatai alapján 2011-ben indult el a Contenta-projekt. A Contenta az egyetemi könyvtár közreműködésével készülő repozitóriumok közös gyűjtőhelye. A különböző típusú digitális szövegek más szoftveres környezetet igényelnek, ezért jelenleg a könyvtár az egyetemi kiadványok (SZTE Egyetemi Kiadványok Repozitóriuma⁸), egyetemtörténeti anyagok (SZTE UnivHistória⁹), és állományvédelmi okból digitalizált folyóiratokat (SZTE Miscellanea¹⁰) a Bodza rendszerben, a szakdolgozatok, doktori disszertációk, és az egyetemen született cikkek kezelését egy E-prints alapú repozitóriummal oldja meg.

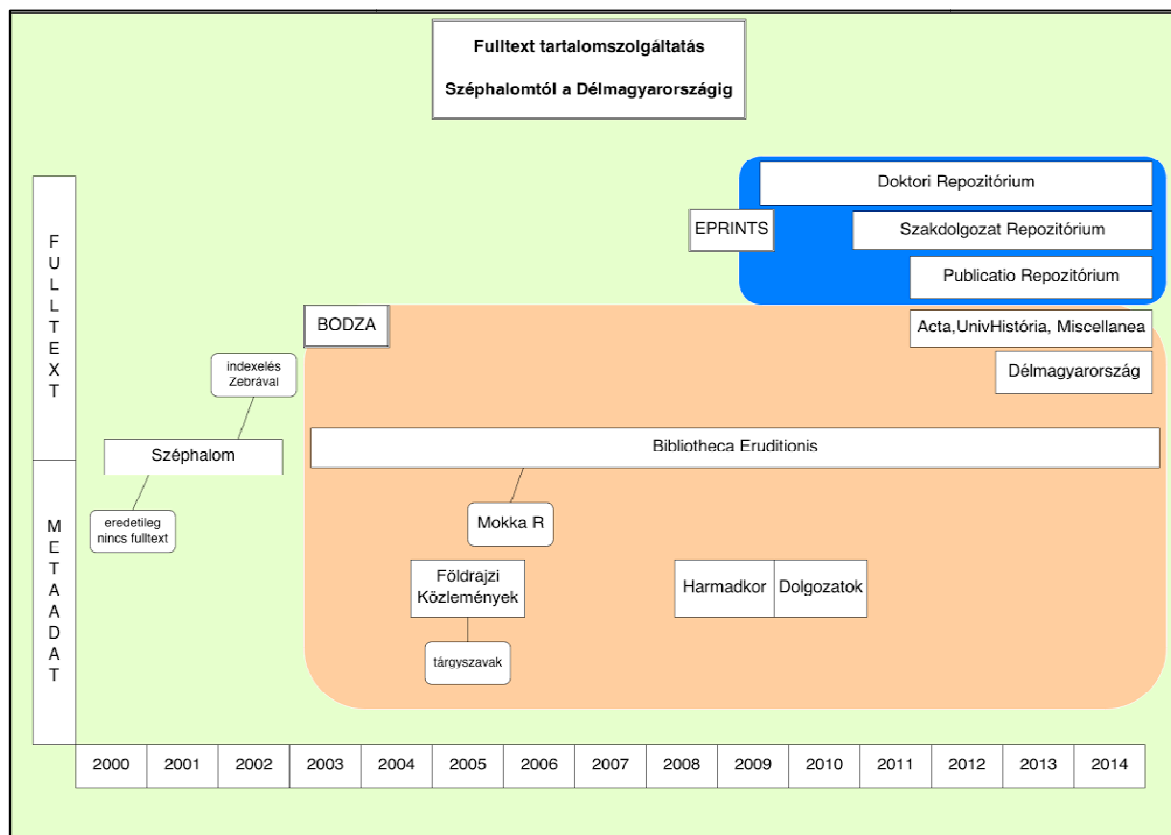
2012 novemberében adták át¹¹ a Contenta repozitóriumot. A 2013-as Networkshop konferencián is készült előadás az elért eredményekről¹².

A digitális Délmagyarország, a DélmagyArchív

A Délmagyarország digitalizált példányainak szolgáltatása, azaz a DélmagyArchív az Egyetemi Könyvtár legújabb, és a feladat nagyságát tekintve

eddig legösszetettebb folyóirat-digitalizálási projektje. A lap fennállásának 100 éves évfordulóján került szóba először a Délmagyarország digitalizálása¹³.

Fontosnak tartom dióhéjban bemutatni a folyóirat történetét, így megvilágítva annak sajtótörténeti jelentőségét. A Délmagyarország Csongrád megye egyik legjelentősebb napilapja, 1910 májusában indult. A ma 104 éves vidéki napilap története szorosan összefügg a 20. század magyar történelmével. Indulásakor már öt különböző jellegű napilap volt a piacon. Ezek közül „a két nagy” a *Szegedi Híradó*, és a *Szegedi Napló* volt, emelet három kisebb jelentőségű lap, a *Szegedi Friss Újság* (korabeli bulvárlap), a *Szeged és Vidéke* (a Nyugat szellemiségéhez hasonló) és a *Friss Hírek* szolgáltatta ki a szegedi lakosság igényeit. Üzletileg kockázatos vállalkozás volt hatodik szereplőként színre lépni, de a Délmagyarország szerkesztőinek és újságíróinak munkája meghozta gyümölcsét. „A Délmagyarország, nem is titkoltnan, a szegedi sajtópiac első helyéért, a vezető pozíció megszerzéséért indult harcba.”¹⁴



1. ábra A szegedi tartalomszolgáltatási projektek összefoglalása

A húszas évek kül- és belpolitikai változásai, valamint a nagy gazdasági világválság alaposan átrendezte a szegedi sajtópalettát. Megjelent két tipikusan 20. századi folyóirat, a szociáldemokrata szellemiségű *Munka*, illetve a revizionista *Szegedi Új Nemzedék*. Az anyagi nehézségek miatt csak a Szegedi Napló, a Délmagyarország, valamint a Szegedi Új Nemzedék maradt talpon, a többi lap csődbe ment. A két világháború közötti időszakban élte a Délmagyarország egyik virágkorát, de mivel ellenzéki lapnak számított, emiatt gyakran érték támadások. 1944-ben a szovjet csapatok bevonulásakor minden szegedi lap megszűnt, csak a Délmagyarország indult újra. Hamarosan a kommunista párt befolyása alá került, később az MKP, majd az MSZMP helyi lapja lett. 1990-ben privatizálták, ma a *Lapcom Kft.* tulajdona.

Az korszakot összefoglaló szegedi sajtótörténet 1855–1945 című ábrán (2. ábra) látható mely hírlapok voltak a Délmagyarország kortársai. A *Munka*¹⁵ és a *Szegedi Ujság*¹⁶ töredékes sorozata szabadon elérhető az SZTE Miscellanea repozitóriumból.

Az előkészületek

A Délmagyarország-projekt eredetileg a szegedi Egyetemi Könyvtár, és a Somogyi-könyvtár közös

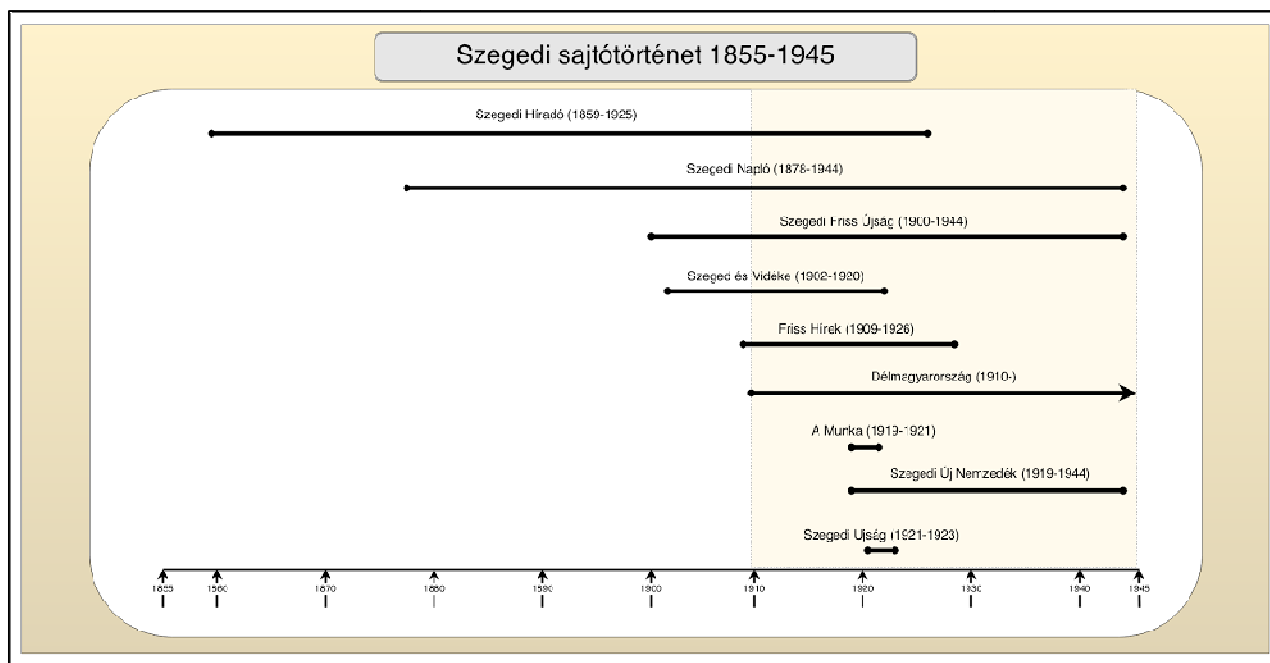
vállalkozása volt, ehhez később csatlakozott a Délmagyarország kiadója is.

Az első állomás, a feldolgozni kívánt kötetek számbavétele, valamint a meglévő könyvtáron belüli, és külső raktárakban tárolt kötetek állapotfelmérése volt.

A digitalizálás több szálon indult el. Az Egyetemi Könyvtár az 1920–1945-ig tartó időszakot, míg ezzel párhuzamosan a Somogyi-könyvtár az 1910–1919 közötti időszakot dolgozta fel. A digitalizálás két különböző eszköztípuson: könyvskenneren, illetve dokumentumskenneren történt.

A digitalizálás folyamatában, a kép elkészülte csak az első szakasz, az utómunkák arányaiban több időt igényelnek. A Délmagyarországot a munkát végző szkennertípusától függetlenül képfájlként digitalizáljuk. Az elkészült fájlokat OCR-ezzük, majd az így elkészült kötegfájlt a PDFToolKit nevű program segítségével bontjuk lapszámra. Ennek a programnak az az előnye, hogy parancssorból futtatható, és jóval gyorsabb mint más PDF kezelő alkalmazások.

Az OCR a sorok ferdeségét a laptükör enyhe forgatásával kompenzálja, ezért a lapokat egyesével körbe kell vágni. Ezzel párhuzamosan történik az egyes lapszámok ellenőrzése is, különös tekintettel az esetleges laphiányokra.



2. ábra Összefoglaló ábra a szegedi sajtótörténetről

A felmerülő hiányokat az *Országos Széchényi Könyvtár* állományából pótoljuk, de sajnos így is maradnak hiányok. Jelenleg egy nyilvános listán keressük a hiányzó részleteket¹⁷.

A Bodza keretrendszerbe való betöltést minden esetben a Bodza fejlesztője végzi el. A folyamat során az egybe tartozó PDF fájlokat különálló oldalakra bontja, majd az egyoldalas PDF-ekből készül egy-egy JPG formátumú képfájl is, mely a webes megjelenítést segíti.

Ezután a szövegfelismerő fájlokból kiemeli a szövegréteget, és ez alapján elkészíti a fulltext indexet. Ezek után a feldolgozás során létrejött XML fájlokat betölti a Bodza keretrendszerbe.

A DélmagyArchív nem csak fulltext kereshető, a napilap cikkei tartalmazó cédlakatalógusok, és elektronikus adatbázisok rekordjainak integrálása is folyamatos.

A szolgáltatás webhelye

Az idei 2014-es Networkshop konferencián *Sándor Ákos* „A Délmagyarország című napilap digitalizálása Szegeden” címmel tartott előadást a projektről.

A szolgáltatás online felületének kialakítása során a kiadónak az volt az álláspontja, hogy egyszerű, könnyen hozzáférhető tartalomra van szükség. Szükség van előre gyártott ajánló jellegű tartalmakra, melyek felkeltik a felhasználó érdeklődését. A könyvtárosok ezzel szemben az összetett, bonyolultabb keresőkifejezés megfogalmazására alkalmas alkalmazást akartak fejleszteni. Az eltérő filozófiát talán úgy lehetne a legegyszerűbben érzékeltetni, hogy a kiadó szerint az olvasó nem tudja, mit keres, míg a könyvtári rendszert használónak van egy előzetes információs igénye.

A felmerült igények miatt kompromisszumos megoldás született. Jelenleg az adatbázis két különböző csatornán érhető el, a www.delmagyar.hu/delmagyarchiv, illetve a <http://dm.ek.szte.hu/> oldalon.

Összegzés

Arra kerestem a választ, hogy milyen folyamatok vezettek oda, hogy 2014 nyarán Szegeden át tudjuk adni a város leggazdagabb múlttal rendelkező napilapjának digitalizált mását, legalábbis a lapok első harmadát. Mindez nem jöhetett volna létre megfelelő szoftveres környezet (Bodza), digitalizáló

célgép (Kodak i660), az idén 105 éves napilap, és a részt vevő intézmények együttműködése nélkül.

A Délmagyarország digitalizálásával az egyetemi könyvtár eddig ismeretlen terepen próbálta ki önmagát. Gyümölcsöző együttműködést sikerült kialakítania a Somogyi-könyvtárral, illetve a Délmagyarország szerkesztőségével. Az egyetemi könyvtár elsősorban az SZTE használói számára nyújt szolgáltatást, korábban nem volt példa arra, hogy az Egyetemi Könyvtár ilyen, klasszikus értelemben véve közművelődési könyvtárra jellemző feladatot vállaljon fel.

Hivatkozások és irodalom

- ¹ LISZTES László – PALOTÁS Gyuláné: Széphalom, 1927–1944: irodalmi és tudományos folyóirat: repertórium. A József Attila Tudományegyetem Könyvtárának kiadványai, Szeged, 1970.
http://digit.bibl.u-szeged.hu/00100/00139/00016/szephalom_rep.pdf
- ² BAKONYI Géza: Tartalomszolgáltatás – egy folyóirat digitális feldolgozása. Előadás, Networkshop, 2000. online:
<http://nws.niif.hu/ncd2000/docs/eloadas/29/index.htm>
- ³ Széphalom. <http://www.bibl.u-szeged.hu/szep/>
- ⁴ SÁNDOR Ákos – HEGYI Ádám: Folyóirat indexelése Zebrával. = TMT, 49. évfolyam (2002) 5. szám online:
http://tmt.omikk.bme.hu/show_news.html?id=647&issue_id=41
- ⁵ HEGYI Ádám: A MOKKA-R (Magyar Országos Közös Katalógus – Régi Nyomtatványok) fejlesztéséről. = Könyvtári Figyelő, 51. évfolyam, (2005) 2. szám online:
<http://epa.oszk.hu/00100/00143/00055/hegyi.html#16>
- ⁶ KEVEHÁZI Katalin: Mi újság a MOKKA háza táján? 3. A régi nyomtatványok és kéziratok adatbázisa. = TMT, 56. évfolyam (2009) 3. szám online:
http://tmt.omikk.bme.hu/show_news.html?id=5095&issue_id=502
- ⁷ PETRÓ Julianna: Huszonöt évvel ezelőtt. = Szegedi Egyetem, 57. (2009) 10., p. [9.]
- ⁸ SZTE Egyetemi Kiadványok Repozitóriuma
<http://opac2.bibl.u-szeged.hu/szteda/acta>
- ⁹ SZTE UnivHistória
<http://opac2.bibl.u-szeged.hu/szteda/univ>
- ¹⁰ SZTE Miscellanea
<http://opac2.bibl.u-szeged.hu/szteda/misc>
- ¹¹ PINTÉR M. Lajos: Három új digitális adatbázist adott át az SZTE Klebelsberg Könyvtár

<http://szegedma.hu/hir/szeged/2012/11/harom-uj-digitalis-adatbazist-adott-at-az-szte-klebensberg-konyvtar.html>

- ¹² KOKAS Károly – SÁNDOR Ákos – ZAWIASA Róbert: Contenta, avagy egy teljes körű digitalizálási program rövid életrajza Szegedről. Networkshop előadás 2013.

<http://videotorium.hu/hu/recordings/details/6152,Contenta>

- ¹³ ÚJSZÁSZI Ilona: Digitalizálják a 100 éves Délmagyarországot.

http://www.delmagyar.hu/szeged_hirek/digitalizaljak_a_100_eves_delmagyarorszagot/2214199/

- ¹⁴ LENGYEL András: „Közkatonái a tollnak...” Vázlatok Szeged sajtótörténetéhez. Bába és Társai Kft. 1999. 494 p. ISBN 963 9144 42 8

- ¹⁵ A Munka: szociáldemokrata politikai napilap.
<http://contenta.ek.szte.hu/misc/100233>

- ¹⁶ Szegedi Ujság.
<http://contenta.ek.szte.hu/misc/100234>

- ¹⁷ Segítségét kérjük! Hiányzó lapszámok jegyzéke.
http://www.delmagyar.hu/delmagyarchiv/segitseget_kerjuk_hianyzo_lapszamok_jegyzeke/2385416/

Beérkezett: 2014. X. 19-én.



Molnár Sándor

a SZTE Klebensberg Könyvtár
digitalizáló és tájékoztató könyvtárosa.
E-mail: sandor.molnar@ek.szte.hu