

Műhely

DélmagyARCHÍV
A DÉLMAGYARORSZÁG archívuma
A szolgáltatás logója

Hírlapok digitalizálása és szolgáltatása a világban és Magyarországon

A Délmagyarország című napilap digitalizálása és szolgáltatása

NAGY Gyula

Bevezető

Két évvel ezelőtt a Szegedi Tudományegyetem (SZTE) Klebelsberg Könyvtárában nagy fába vágtuk a fejszénket: elkezdük az 1910-ben alapított *Délmagyarország* című napilap digitalizálását.¹ A vállalkozásról több országos szakmai rendezvényen beszámoltunk^{2,3} már, és itt az ideje, hogy írásban is összefoglaljuk a lap digitalizálásával kapcsolatos legfontosabb eredményeket, tanulságokat és buktatókat.* A digitalizálási projekt először ijesztőnek tűnt,

ugyanakkor kecsesgőzösen izgalmas is volt. Mára az ijedtség elmúlt, de a kihívás és izgalom megmaradt, és jelen lesz egészen addig, amíg az utolsó lapszámot be nem tesszük a szkennerbe. A feladat izgalmas része és érdekes vetülete egyrészt az ilyen típusú újságok alapvető tulajdonságából következik, mégpedig, hogy a mindennapok első vonalából közölnek tudósításokat. Így a történelem olyan szintjei és részletei maradnak fent a megsárgult lapokon, amelyek máshol nem érhetők el, és ezzel a kor eredeti és hiteles rajzát őrzik meg az utódok számára. A másik része, ami

* Az írás eredetileg „Az év fiatal könyvtárosa” pályázatra készült, a most közölt tanulmány annak cikkésített, rövidített változata.

izgalmassá teszi a feladatot, a projekt nagysága, a digitalizálandó oldalak mennyisége.

A Délmagyarország esetében az előzetes felmérések során nagyjából 300 és 500 ezer oldal közöttire becsültük a teljes terjedelmet, amelyről jelenleg (körülbélül a munka felénél) azt gondoljuk, hogy a végeredmény inkább az alsó becsüléshez lesz majd közelebb. A vállalkozásunkról megjelenő első, 2011-es híradás után egy hozzászóló viccesen csak annyit jegyzett meg, hogy „Hurrá, valakinek megvan a nyugdíjig az álommelója!”⁴ Akkor még nem tudtuk, milyen ütemben lehet majd haladni, ma már látjuk, valószínűleg nem lesz igaza a kommentelőnek. Nagyjából két év alatt elkészültünk a munka felével, úgy, hogy közben számos egyéb digitalizálási és könyvtárosi feladatot is el kellett látnunk. Ez csak úgy sikerülhetett, hogy több helyről is jelentős támogatást kaptunk: a munka sikere két könyvtár együttműködésének köszönhető, ugyanis összefogtunk a Somogyi-könyvtárral, melynek munkatársai a mai napig sokat segítenek a projekttel kapcsolatos munkákban. A másik pótolhatatlan segítséget saját könyvtárunk munkatársaitól kaptuk: az intézmény dolgozóinak legalább a fele valamilyen formában közreműködik a munkában.

Felmerülhet a kérdés, miért van szükség ennyi ember munkájára egy napilap digitalizálásához (és egyáltalán hogyan fér ennyi könyvtáros a szkennel mellé). A projekt jóval többről szól, mint újságoldalak szkenneléséről: mindenkinek akad testreszabott feladat, amellyel előbbre mozdítható a projekt ügye.

Az út során, amíg eljutottunk a mostani állapothoz, alaposan kitanultuk, hogyan érdemes hírlapokat digitalizálni, feldolgozni és szolgáltatni. Bízunk benne, hogy a digitalizálás máshol is a mindennapi könyvtári gyakorlat része lesz, és hogy az általunk kidolgozott megoldások hasznos tapasztalatokkal szolgálnak a többi könyvtárnak is. Igaz, a hazai könyvtárak egyelőre még nem állnak túl jól a hírlapok digitalizálása terén, de kétségtelen, hogy ebben az évben már komoly kezdeményezések történtek.

Mi az a hírlap?

A *Magyar Nagylexikon* szócikke így jellemzi a hírlapokat:

„Hírlap vagy újság: az időszaki sajtó egyik fő típusa. Leggyakoribb megjelenési formája a napilap. A legaktuálisabb napi eseményekről nyújt széles körű tájékoztatást, olvasóinak száma többnyire meghaladja a folyóiratokét, ezektől eltérően részletesebb, tudományos jellegű értekezéseket nem közöl. Előfutárai között a kéziratú újságokat és nyomtatott újságleveleket említik. Az első rendszeresen megjelenő hírlapok Németországban jelentek meg (*Relation: Strassburg, 1605; Aviso: Wolfenbüttel, 1609*). A magyarországi latin és német nyelvű hírlapok sorát az 1705-ben meginduló *Mercurius Hungaricus (Veridicus)*, II. Rákóczi Ferenc hadi lapja nyitotta meg. Az első magyar nyelvű hírlap Pozsonyban jelent meg *Magyar Hírmondó* címmel (1780). A 18. század végén és a 19. század első felében kibontakozó magyar hírlapirodalom a kedvezőtlen politikai, társadalmi és cenzúráviszonyok miatt csak lassan válhatott a politikai élet szerves részévé. Magyarország első valódi politikai hírlapja a Kossuth Lajos által 1841-ben megindított *Pesti Hírlap* volt. Az 1867-es kiegyezést követően létrejöttek az első nagy üzleti lapvállalatok (a Légrády testvérek *Pesti Hírlapja, 1878; Rákosi Jenő Budapesti Hírlapja, 1881*). A századforduló táján megjelentek az első bulvárlapok is, olykor több százszáz példányban”.⁵

A szócikkben kicsit összemosódik a hírlap és a napilap fogalma, érdemes tehát más forrást is megnéznünk, hogy kiderüljön miben áll a különbség a két rokon kifejezés között. A *Könyvtárosok kézikönyve* így ír:

„A hírlapok általában rövid időközönként (naponta, kétnaponta, hetenként) megjelenő kiadványok, amelyek fő célja az aktuális politikai társadalmi eseményekről való tájékoztatás, a mindennapi életben széles körű érdeklődésre számot tartó, gyors információk közlése.”⁶

A napilapok és a hírlapok között húzódó fő különbség tehát a periodicitásban ragadható meg. Azt is látnunk kell, hogy a hírlap fogalma ma-

gában foglalja a napilap fogalmát. Minden napilap hírlap, de nem minden hírlap napilap, hiszen amelyek nem naponta jelennek meg, de híreket közölnek, azok megfelelnek a hírlap kategória ismérveinek, de nem elégítik ki a napilapokét.

Néhány további, a hírlapokra jellemző ismérv:
 • híreket, közéleti eseményeket, bulvárt tartalmaznak
 • formai megjelenésüket tekintve általában nagyméretűek
 • egy-egy oldalon rövidebb-hosszabb cikkeket, híreket találunk (az egész oldalas vezércikktől a néhány mondatos tudósításig)
 • sokszor naponta jelennek meg
 • ugyanakkor nem kizárólagosan, hiszen vannak heti, kétheti, havi megjelenésűek is (például egy-egy városi vagy falusi újság esetében inkább ez a tipikus)
 • az előző pontban említettek helytörténeti jelentősége kiemelten fontos
 • sok esetben családtörténeti jelentőségük is lehet (születési, házassági és gyászjelentéseket közölnek, feltárják a személyek politikai vonatkozásait és tetteit, ill. családokhoz fűződő vállalkozások reklámjai, apróhirdetések is jellemzik, amelyekből nem várt részletességű adatok tudhatók meg a kutató számára).

A „tegnapi újság”

Elsőre talán nem egyértelmű, de ha jobban belegondolunk, a „tegnapi újság” tartalma is foglalkoztatni szokta az embereket. A napilapokban gyakran találkozhatunk olyan rovatokkal, amelyekben visszatekintenek öt, tíz, húsz, ötven, nyolcvan vagy száz évvel korábbi eseményekre, felidézve, hogy egy adott napon mi volt hírértékű információ (általában saját lapjuk múltjából ragadnak ki híreket és cikkeket). Az embereket érdekli a múlt, különösen a múlt hétköznapi vetülete. Ezt az igényt elégíti ki pl. a népszerű internetes portál, az Index.hu 2014-ben indult rovata, *A nagy háború*⁷ is, amelyben az 1896-ban alapított *Esti Újság* 1914. évi számaiból közölnek napi válogatást az első világháború százéves jubileuma alkalmából. Bár *Buzinkay Géza* sajtótörténész bulvárlapként utalt e lapra,⁸ az *Esti Újság* sem külsejében, sem tartalmában nem hasonlít a mai értelemben vett bulvárlapokra: a dolgozó néprétegeknek szánt könnyen emészthető, politikai, közéleti sajtótermék volt, mely négy hasámban tördelt tipográfiájával is a

hírlapi hagyományokat folytatta.

Néhány hazai könyvtár (pl. az OSZK,⁹ a FSZEK vagy a szegedi Somogyi-könyvtár,¹⁰ sőt számos magáncég) kínálatában szerepel az ún. születésnapi újság („születésnapi csomag”) szolgáltatás, amely azon túl, hogy bevételi forrást jelent a szolgáltató intézmény számára, valós olvasói igényt elégít ki, ezért pozitívként kell megemlítenünk. A szolgáltatás lényege, hogy régi újságokból digitalizált és kinyomtatott válogatásokat készítenek arról a napról, amikor a csomaggal megajándékozni kívánt született. Ha nem lenne igény erre a szolgáltatásra, akkor bizonyosan nem nyújtana ilyen szolgáltatást ennyi piaci szereplő, elmondhatjuk tehát, hogy a „tegnapi újság” ajándéknak sem utolsó.

Bizonyos regényekben és filmekben is visszatérő motívumként jelennek meg az újságok. Lelkes filmkedvelők ezekről képes válogatást¹¹ és videót¹² is összeállítottak.

A hírlapok digitalizálásának kérdésköre nem csak a könyvtárosokat foglalkoztatja. Például a *Nyelv és Tudomány* portál egyik cikke a digitalizálásra szánt anyagok kiválasztása kapcsán az emlékezet és a nyelv utólagos befolyásolásának problémáját boncolja:¹³ jogosan merülhet fel egy-egy hírlap-digitalizálási projekt kapcsán annak megválaszolása, miért éppen az adott újság kerül digitális megőrzésre.

A hírlap-digitalizálás a szakirodalom tükrében

A hírlapok digitalizálása jelentősen különbözik más dokumentumtípusok digitalizálásától, illetve szolgáltatásától. Nem véletlen, hogy általában külön erre a célra tervezett adatbázisokban szolgáltatják őket.

A hírlapok digitalizálásának kérdéseiről először a különféle műhelyek weboldalain tájékozódtam, majd megnéztem, hogyan jelent meg e téma a könyvtári szakirodalomban. Példáim kiragadottak, szubjektív válogatást tükröznek, de így is jól érzékeltetik a hírlap-digitalizálás jellegzetességeit.

Egy összefoglaló tanulmány (*The unique challenges of newspaper digitization*) részletesen kifejti,

miért igényelnek speciális bánásmódot a hírlapok. Elsőként a nagy és nem egységes lapméretet emelik ki, amelyet sok esetben csak nagyméretű, ipari eszközökkel lehet beszkenneálni. A nagyobb fizikai lapméret értelemszerűen nagyobb, megabájtokban kifejezhető képméretet eredményez. A cikk szerint egy tipikus könyvoldalon átlagosan 400–600 szó fordul elő, egy átlagos újságon 4000 és 6000 közé tehető. Ez természetesen tükröződik a felismerési és indexelési időkben és a nagyobb adatbázisméretben is. Megemlítik az újságok speciális szerkezeti és tipográfiai felépítését: a fejléceket, a rövidebb, hosszabb cikkeket, a hasábtördelést, az illusztrációkat, képeket, reklámokat. Kiemelik a METS/ALTO szabvány hasznosságát, amellyel jól kezelhető ez az egyedi szerkezeti struktúra. Írnak a karakterfelismerési hibák előfordulásáról és lehetséges javítási módokról, megemlítve a felhasználók általi manuális OCR-javítás lehetőségét is, amely csak akkor lehetséges, ha erre a szolgáltatást biztosító szoftvermotor is alkalmas.¹⁴

Egy másik weboldal (*Best practices for newspaper digitalization*) alternatív szkennelési forrásra, a mikrofilmekre hívja fel a figyelmet. A skandináv országokban már évekkel korábban készült olyan átfogó digitalizálási projekt, amelyben a mikrofilm-digitalizálás és az OCR technológia technikai megvalósíthatóságát és gazdaságosságát vizsgálták. A mikrofilmek szkennelése jól automatizálható, ezáltal egyszerűbb és olcsóbb maga a digitalizálás. Az oldalon hasznos tanácsokat adnak a fájlok elnevezésével, minőségével és felbontásával kapcsolatban is. A használt fájlnevsablon azért nagyon fontos, mert a fájlneveknek több kritériumnak is meg kell felelniük. Többek között biztosítaniuk kell a könnyű visszakereshetőséget és az oldalak eredeti sorrendjét is.¹⁵

A Kongresszusi Könyvtár nagyszabású újságdigitalizálási projektje, a National Digital Newspaper Program honlapján (<http://www.loc.gov/ndnp>) számos segédanyag érhető el. Ilyen például a program 76 oldalas gyakorlati alkalmazását ismertető dokumentuma (The National Digital Newspaper Program (NDNP) Technical Guidelines for Applicants.¹⁶) A honlap egyik

oldalán (National Digital Newspaper Program Guidelines & Resources) a témával foglalkozó aktuális segédleteket és cikkeket gyűjtötték össze.¹⁷

A British Library is régóta foglalkozik a hírlapok digitalizálásával. Az általuk végigjárt utat mutatja be *Ed King* írása a The British Library newspaper collection and the British Newspaper Archive¹⁸-ról.

Allison Quam egy helyi kezdeményezésről számol be: hogyan sikerült Minnesotában széles körű összefogással digitalizálni, használhatóvá tenni a legfontosabb helyi lapok régebbi évfolyamait. Quam cikke¹⁹ azért is érdekes, mert a projektet a minnesotai egyetemi könyvtár kezdeményezte, és hozzánk hasonlóan, nekik is sikerült megszervezniük az együttműködést a projekt köré, amely lehetővé tette a közös munkát, ahogy ez volt a mi célunk is Szegeden.

A hazai szakirodalomban kevés, kifejezetten a hírlapok digitalizálásával foglalkozó írást találhatunk, ami valószínűleg nem véletlen, hiszen, mint ahogyan a *Hazai kezdeményezések* című alfejezetben látni fogjuk, jelenleg Magyarországon csak két komolyabb méretű, magyar hírlapokat tartalmazó digitális gyűjtemény üzemel. Hírlap-digitalizálási próbálkozások voltak, sőt hírlap-digitalizálás is folyt/folyik hazánkban, azonban ezen kezdeményezések közül még csak kevés ért révbe. Nincsenek azonban kétségeim, hogy ez hamarosan változni fog, ismerteve azokat a munkákat, amelyek bizonyos műhelyekben jelenleg is zajlanak.²⁰ Utalnék még egy összefoglaló cikkre (Veroszta: *Kerekasztal-beszélgetés a digitalizációról, az emlékmentésről és az ezzel együtt járó szelekcióról*²¹), amelyben bár csak részlegesen érintik a hírlap-digitalizálás nehézségeit, de különböző aspektusokból jól körüljárják a digitalizálás alapkérdéseit. Néhány érdekes elképzelés mindenesetre a napilapokkal kapcsolatban is elhangzott a beszélgetés során, bizonyítékként a fenti jóslatomnak, hogy hamarosan lesznek hazai digitális hírlaparchívumaink.

Külföldi hírlaparchívumok

• Kulcsszavak

Ha külföldi hírlaparchívumokban szeretnénk tájékozódni, jó, ha ismerjük az angol nyelvű terminológiát és a kapcsolódó fogalmakat.

Újság: newspaper, news-sheet, news, journal, print, tidings, *hírlap*: newspaper, news-sheet, *napilap*: daily newspaper, daily paper, daily, gazette, *újság-archívum*, *adatbázis*: newspaper archive, newspaper database, newspaper collection. *Egyéb kapcsolódó kifejezések*: newspaper digitization, digital repository.

• Gyűjtőoldalak

A kulcsszavak tisztázása után már célzottabban tudunk elindulni a világhálón a témát kutatva. Kiindulópontként először mindenképpen a jelentősebb hírlaparchívumokat listázó gyűjtőoldalak áttekintése javallott.

– Wikipedia: List of online newspaper archives (http://en.wikipedia.org/wiki/Wikipedia:List_of_online_newspaper_archives)

Az első gyűjtőoldal, amivel érdemes foglalkoznunk a Wikipédián összegyűjtött hírlaparchívumok listája. Itt szinte minden jelentős archívum szerepel, sőt azt kell mondanunk, hogy inkább a bőség zavara jellemző. Szerencsére, a lista országok szerint bontott, ami megkönnyíti az eligazodást. Elindulásként mindenképpen ajánlott ennek a gyűjteménynek a tanulmányozása, magam is ezt használtam elsődleges tájékoztató forrásként, a bemutatott archívumok egy részéhez innen jutottam el.

– Historic Newspapers Online (http://genealogy.about.com/od/newspapers/tp/newspapers_online.htm)

A *Kimberly Powell* által összeállított gyűjtőoldal, az About.com Genealogy aloldalához tartozik. Elsősorban családtörténeti megközelítésű, de ez mit sem von le az értékéből, jóval szűkebb, ugyanakkor minőségi válogatásról van szó. Akit zavarba ejt az előző forrás összetettsége, annak itt érdemes próbálkoznia.

– International Coalition on Newspapers (<http://icon.crl.edu/digitization.php>)

A harmadik gyűjtőoldal, az ICON az első forráshoz hasonlóan minél több adatbázis listázását tűzte ki célul: olyan kataszter akar lenni, amely nyilvántartja az egész világon zajló újságdigitalizációs projekteket és az elkészült archívumokat. Fontos megemlíteni a könyvtári kapcsolatot is, ugyanis észak-amerikai és kanadai könyvtárak állnak a kezdeményezés mögött.

– Elephind.com: Search the world's historic newspaper archives (<http://www.elephind.com>)

Végül a gyűjtőoldalokhoz kapcsolódóan egy keresőt említhetünk még, ahol közös felületen teljes szövegben tudunk keresni Ausztrália, Új-Zéland, Szingapúr, Mexikó és az Amerikai Egyesült Államok legfontosabb szabadon hozzáférhető archívumaiban. Mindenképpen hiánypótló szolgáltatásról van szó, a honlapon található információ szerint a közös keresőt folyamatosan bővítik.

• Nemzetközi és nemzeti archívumok

A gyűjtőoldalak böngészése során jól látható, hogy kétféle szemlélet jellemzi az archívumok működését. A non-profit megközelítés mögött általában könyvtárak állnak, a for-profit alapúak esetében profitorientált cégek biztosítják a megvalósítást és a szolgáltatást. Könyvtárosként számunkra az első típus az érdekesebb, a második csoporttal csak érintőlegesen foglalkozom, bemutatva néhány nagyobb fizetős adatbázist (többségükben a családtörténeti kutatásokat támogatják, és ezen a területen bevett gyakorlat, hogy az érdeklődő genealógusoknak fizetniük kell a használatért).

• Nemzetközi adatbázisok

– Gale Digital Collections

Olyan cégcsoportról van szó, amely elsősorban könyvtáraknak és más intézményi partnereknek kínálja előfizetéses adatbázisait, a magánfelhasználókra kevésbé koncentrálnak. Adatbázisai köre szerteágazó, hírlaparchívum csomagjaik közül néhány kiragadott példa: 19th Century

U.S. Newspapers, British Newspapers 1600–1950, Daily Mail Historical Archive 1896–2004, The Times Digital Archive 1785–2006. Fizetős adatbázis, a <http://gdc.gale.com> címen érhető el.

– Newspaperarchive.com

Szerényen a világ legnagyobb gyűjteményeként aposztrofálják magukat, és az adatok valóban impozánsak: több mint 130 millió digitalizált oldal, mintegy 400 év újságjai találhatóak benne az Egyesült Államokon kívül tíz további országból is. 1999-es indulás óta kétszázezer előfizetőt gyűjtöttek. Többféle böngészési lehetőséget biztosítanak: lehet időben, földrajzi térben, újság-cím szerint mozogni, és teljes szövegű keresésre is lehetőség van.

Fizetős adatbázis, a <http://newspaperarchive.com> címen érhető el.

– Newspapers.com

Az előző szolgáltatáshoz hasonló a kínálat, de itt előfizetés nélkül is meg tudjuk nézni, hogyan működik a rendszer. Terjedelmét tekintve körülbelül az előző adatbázis felére tehető a digitalizált oldalak száma (látogatásom idején 66 millió oldalnyi tartalom volt hozzáférhető), de folyamatosan gyarapszik a gyűjtemény. A 18–21. századot átfogóan mintegy 2800 cím található, főként az USA tagállamaira koncentrálnak, de szerepelnek Kanada, Anglia, Írország és Panama újságjai is.

Kiemelendők még a Web 2.0-ás eszközök: amelyben a regisztrált felhasználók tudnak újságkivágásokat készíteni, a rendszer listázza az egyes újságokhoz köthető felhasználókat és széles körű megosztási lehetőségeket is biztosít.

Ingyenes/fizetős adatbázis, a <http://www.newspapers.com> címen érhető el.

– Genealogybank.com

Családtörténeti profilú előfizetéses oldal, több mint 6500 újság szövegében lehet keresni 1690-től kezdve napjainkig. A hírlapokon kívül sok mást is kínál: többek között születési és házassági kivonatokban, gyászjelentésekben, utaslistákban és katonai összeírásokban is lehet tájékozódni. Csak amerikai újságok szerepelnek

az adatbázisban.

Fizetős adatbázis, a <http://www.genealogybank.com> címen érhető el.

– Google News Archive

A Google a Google Books könyvdigitalizálási sikerein felbuzdulva 2008-ban indította el a hírlap-digitalizálási projektjét, amelynek során mintegy 60 millió oldalt szkenneltek be, tettek hozzáférhetővé körülbelül 2000 újságból. Sajnos 2011 májusában a projektet leállították a felmerülő nehézségek miatt.²²

Főként az USA, az Egyesült Királyság és Kanada képviselteti magát angol és francia nyelvű hírlapokkal. Az adatbázis szabadon hozzáférhető, a Google-től elvárható módon a keresési funkció kiválóan működik, a böngészés már kevésbé van kidolgozva. Látszik, hogy befagyasztották a szolgáltatást, pedig az archívumból jóval többet is ki lehetne hozni egy kis fejlesztéssel, a feltöltött tartalom ingyenesen elérhető és használható.

Ingyenes adatbázis, a <http://news.google.com/newspapers> címen érhető el.

Nemzeti adatbázisok

👁 Amerikai Egyesült Államok

– Chronicling America Historic American Newspapers

Az Amerikai Egyesült Államok legnagyobb szabadon hozzáférhető archívuma, amelyet a Kongresszusi Könyvtárhoz köthetően a National Digital Newspaper Program (NDNP) keretében alakítottak ki és töltöttek fel tartalommal. Jelenleg körülbelül 8 millió oldalnyi teljes szövegű anyagot tartalmaz. Az archívum jól strukturált, gyorsan, könnyen használható, sokféle hasznos módszertani segédanyagot is tartalmaz (ezekről bővebben a *Hírlap-digitalizálás a szakirodalom tükrében* című alfejezetben). A kezdőoldalon a klasszikus „tegnapi újság” fogad bennünket, kilistázva a napra pontosan száz évvel ezelőtt megjelent hírlapok összes olyan számát, amely hozzáférhető az adatbázisban. A kereső és böngésző funkciók jól használhatóak, az újságoldalak egy saját megjelenítőben érhetőek el, amely minden felmerülő igényt kielégít.

Végül a Stanford Egyetemen készült térképes megjelenítést adó adatbázisra (*Data visualization: Journalism's voyage West*) hívnám fel a figyelmet:²³ az archívum kiválóan használható böngészésére, lehetővé teszi mind az időbeli, mind a térbeli navigálást.

Ingyenes adatbázis, a <http://chroniclingamerica.loc.gov> címen érhető el.

– NYS Historic Newspapers

Nagyon hasonló a Chronicling Americához, mivel ugyanazt a keretrendszert használja. A Northern New York Library Networkhöz köthető, amely New York államhoz (és nem New York Cityhez) tartozik, központjuk a kanadai határ mellett található. Az adatbázisban körülbelül 300 újság három millió oldalnyi terjedelemben kereshető vissza.

Ingyenes adatbázis, a <http://nyshistoricnewspapers.org> címen érhető el.

Kanada

© Multicultural Canada

Kanadában nagyrészt tartományi szintű archívumok vannak, amelyeket terjedelmi okok miatt egyenként nincs módomban felsorolni. (Egyébként az amerikai modellhez hasonlóan, ott nem említettem ezt, csak kiragadtam a legnagyobb gyűjteményeket.) Azonban Kanada élen jár a hírlapok digitalizálásában, ezért mindenképpen szerepeltetni szerettem volna, ezért választottam egy több tartományra kiterjedő gyűjteményt. Aki más kanadai archívumot szeretne böngészni, a korábban ismertetett gyűjtőoldalakon szép számmal találhat.

Ingyenes adatbázis, a http://multiculturalcanada.ca/contentdm_search címen érhető el.

Egyesült Királyság

© The British Newspaper Archive

Jól használható, jól „összerakott” adatbázis, nagy hátránya, hogy csak díjfizetés ellenében használható, és az előfizetési díjak viszonylag magasak. Van kedvezményes opció is: két napra szóló 100 oldalas hozzáférés 7 font (kb. 2600 forint), de az éves, korlátlan hozzáférés már 80 fontba (30.000 forint) kerül.²⁴ Ez abból adódik,

hogy a British Library ugyan közreműködött az archívum létrehozásában, azonban a digitalizálást és a szolgáltatást is a DC Thomson cégcsoport részét képező Brightsolid végzi.

Az adatbázis jelenleg 250 címet és 8 millió oldalt tartalmaz (tíz éven belül 40 millió oldalra tervezik bővíteni), a szokásos funkciók mind rendelkezésre állnak: teljes szövegű keresés, böngészés dátum, földrajzi hely és cím szerint. Regisztráció nélkül még az ingyenes példaoldalakhoz sem férünk hozzá. Bár a könyvtár olvasótermeiben ingyenesen használható, ezt a szigorú hozzáférési modellt negatív példaként kell megemlítenünk. Szerencsére a brit gyakorlat egyedülálló, minden más esetben azt tapasztaljuk, hogy ha könyvtárak állnak egy ilyen szolgáltatás mögött, az adatbázis szabadon hozzáférhető a világhálón.

Fizetős adatbázis, a <http://www.britishnewspaperarchive.co.uk> címen érhető el.

Ausztrália

© Trove Digitised Newspapers and more

A briteknél szerzett negatív tapasztalat után az ausztrálok hozzáállásában az ellenkező véglet figyelhető meg, amit ambiciózus tervükben, az *Australian Newspaper Digitisation Program*-ban²⁵ rögzítettek. Céljuk, hogy teljesen ingyenesen, online hozzáférhetővé tegyenek minden, az 1950-es évek közepéig megjelent újságot, párhuzamosan egy másik kezdeményezéssel, az *Australian Newspaper Plan*-nel²⁶, amelyben valamennyi Ausztráliában kiadott újság összegyűjtését és megőrzését tűzték ki célul, támogatandó a jövőbeli digitalizálási programokat. Jelenleg körülbelül 650 cím és mintegy 13 millió oldal található az archívumban.

Az adatbázis jól kereshető a szokásos böngészési lehetőségek szerint. Érdekessége, hogy az összes vizsgált archívum között itt találtam a legfejlettebb Web 2.0-ás eszközöket, és próbálnak élni a crowdsourcingban (online, sokszereplős, feladatok kiosztására épülő problémamegoldó és termelő modell) rejlő lehetőségekkel is. Ezt a legutóbbi aktivitások monitorozása alapján sikerült is elérnie az oldalnak. A regisztrált felhasználók javítani tudják az automatikus ka-

rakterfelismerés hibáit, ezáltal teljesen hibátlan szövegátiratok kerülhetnek be a rendszerbe, sőt címkézni és kommentálni is tudják az egyes cikkeket. A felhasználó-kezelés nyilvántartja a felhasználók tevékenységét az egyes oldalakon, így lekérhetőek a felhasználókhöz köthető aktivitások, és például toplistákat lehet generálni a legaktívabbakról. Hasonló funkciókat a többi archívumnak is érdemes volna beépíteni, mert a használok fokozottabb bevonásának van jövője. Természetesen az alapvető és már-már kötelező közösségi funkciókat szintén tudja az oldal: pl. különböző (akár cikk szintű) megosztási lehetőségek, fórum.

Ingyenes adatbázis, a <http://trove.nla.gov.au/newspaper> címen érhető el.

Új-Zéland

© Papers Past

Az Új-Zélandi Nemzeti Könyvtárhoz kapcsolódó szolgáltatás. Az adatbázisban az 1839–1945 közötti évekből 83 cím és kb. 3,5 millió oldal terjedelemben található. A rendszer ismeri a szokásos böngészési funkciókat (dátum, régió, cím) és a teljes szövegű keresést. Megjegyzendő, hogy cikk, illetve hasáb szintű feldolgozást alkalmaznak, ezt a METS/ALTO szabványok teszik lehetővé. Ez a tulajdonság megkönnyíti az egyes cikkek analitikus feldolgozását, ami nagy segítség a visszakeresésnél is. Az anyagokat fekete-fehérben szkennelték, ami bizonyos esetekben hátrányt jelent.

Ingyenes adatbázis, a <http://paperspast.natlib.govt.nz> címen érhető el.

Karib-szigetek

© Caribbean Newspaper Digital Library

A Karib-térség országainak együttműködésére épülő digitális gyűjtemény, amelyben több ország hírlapjai hozzáférhetőek. A sok közreműködő miatt a képek minősége nem egységes.

Ingyenes adatbázis, a <http://dloc.com/cndl> címen érhető el.

Izrael

© Historical Jewish Press

Az Izraeli Nemzeti Könyvtár adatbázisában 46

újságcím, 1,2 millió oldal terjedelemben érhető el az 1843–1987 közötti évekből. Nemcsak az Izraelben kiadott hírlapokat gyűjti, hanem a zsidó diaszpóra lapjait is. A szokásos navigációs funkciók megvannak, a megjelenítő felület pedig különösen kényelmes, jól sikerült fejlesztés. Az egyes oldalak itt is cikkekre vannak bontva.

Ingyenes adatbázis, a <http://web.nli.org.il/sites/JPress/English> címen érhető el.

Európai Unió

© Europeana Newspapers

Látogatásom idején még a prototípus működött, azonban már így is tetemes mennyiségű anyag volt hozzáférhető, és a keresési, böngészési funkciók is rendben voltak. Az Europeana modelljének megfelelően az anyagokat ebben az esetben is az egyes országok együttműködő partnerei biztosítják a szolgáltatás számára. Mintegy 2200 hírlap található az adatbázisban, a legnagyobb beszállító Spanyolország és Izland.

Ingyenes adatbázis, a <http://www.theeuropeanlibrary.org/tel4/newspapers> címen érhető el.

Luxemburg

© Eluxemburgensia

Jól felépített szolgáltatásról van szó, látszik, hogy nem sajnálták rá a pénzt. A cikk szintű feldolgozást használják a METS/ALTO szabványokkal együtt. Külön ablakos megjelenítője van, ami rendkívül funkciógazdag, még a cikk szintű linkelhetőséget is biztosítja.

Ingyenes adatbázis, a <http://www.eluxemburgensia.lu> címen érhető el.

Svájc

© Swiss Press Online

Az ország hivatalos nyelvei szerint csoportosítva található a rendszerben lévő hírlapok. Az Olive cég keretrendszerét használja az oldalak megjelenítésére.

Ingyenes adatbázis, a <http://newspaper.archives.rero.ch> címen érhető el.

Izland

© Timarit.is

Izland, Grönland és a Feröer-szigetek közös hírlaparchívuma 888 címmel és 4,5 millió oldalnyi digitalizált tartalommal (imponálóak az adatbázis mennyiségi jellemzői a lakosság nagyságához képest). Szinte az összes eddig megjelent újságot tartalmazza, a honlapon elérhető leírás szerint. Csak cím szerinti böngészést kínálnak, és a keresési funkció is lehetne bővebb, a megjelenítés átlagos.

Ingyenes adatbázis, a <http://timarit.is> címen érhető el.

Finnország

© Historical Newspaper Library

A Finn Nemzeti Könyvtár adatbázisa 1,7 millió oldallal az 1771 és 1900 között megjelent hírlaptartalommal. Az oldal angolul is elérhető, a cím szerinti böngészés, az összetett, illetve a teljes szövegű keresés biztosított. A digitalizált tartalom megjelenítése nem különösebben kifinomult, de használható, lehetőség van PDF-formátumú letöltésre is.

Ingyenes adatbázis, a <http://digi.lib.helsinki.fi/sanomalehti> címen érhető el.

Franciaország

© Gallica

A franciák mindig is élen jártak a digitalizálásban, így nem meglepő, hogy a hírlapok területről sok minden hozzáférhető. Sajnos az oldalnak csak bizonyos részei érhetőek el angolul, és az is nehezíti a használatot, hogy az újságokat nem kezelik külön (pedig szerencsésebb volna egy ilyen megoldás), így a Gallica általánosabb eszközkészletével kell beérnünk.

Ingyenes adatbázis, a <http://gallica.bnf.fr/html/presse-et-revues/les-principaux-quotidiens> címen érhető el.

Németország

© Zeitungsinformationssystem (ZEFYS)

A Berlini Állami Könyvtárhoz köthető adatbázis. Németországban a tengerentúli példákhoz hasonlóan inkább a tartományi szintű digitalizációs projektek a jellemzőek. Az oldal a német

mellett angolul is használható. A ZEFYS 113 újság számait tartalmazza 1739-től egészen 1946-ig; dátum és a cím szerinti böngészést kínál. Az adatbázis nemcsak digitalizált újságokat tartalmaz, hanem információszolgáltató rendszerként is működik, valamelyest megnehezítve a digitalizált tartalomhoz való hozzáférést, így például a teljes szövegű keresést.

Ingyenes adatbázis, a <http://zefys.staatsbibliothek-berlin.de> címen érhető el.

© DigiPress

A Bajor Állami Könyvtár sajtóarchívuma. Bár béta verzióként hirdetik magukat, de komoly probléma nem tapasztalható az oldal működésében. 28 hírlapot szolgáltatnak, a dátum és cím szerinti böngészés lehetősége adott, ahogy a teljes szövegű keresés is. A kép alapú megjelenítő csak a legfontosabb alapfunkciókat ismeri: nagyítás, lapozás, PDF letöltése.

Ingyenes adatbázis, a <http://digipress.digitale-sammlungen.de> címen érhető el.

Ausztria

© Austrian Newspapers Online (ANNO)

Az Osztrák Nemzeti Könyvtár komoly digitalizáló tevékenységet folytat, még magyar vonatkozású címeket is beszkeneltek és hozzáférhetővé tettek (pl. Der Ungar, Pester Lloyd, Erdélyi Híradó). Az adatbázis körülbelül 500 címet és 10 millió oldalt tartalmaz jelenleg.

Csak német nyelvű használatot kínál, a keresés és a böngészés kicsit nehézkes, azonban a tartalom sokrétű és értékes (nem csak a magyar vonatkozások miatt). A megjelenítés kép alapú, de ettől függetlenül van letölthető PDF-változat is.

Ingyenes adatbázis, a <http://anno.onb.ac.at> címen érhető el.

Hazai kezdeményezések

• Elektronikus Periodika Archívum (EPA)

Ha Magyarországon digitális állományok szolgáltatásáról beszélünk, megkerülhetetlenek a Magyar Elektronikus Könyvtár (MEK) és a hoz-

zá kapcsolódó adatbázisok, hiszen velük indulnak az ilyesfajta kezdeményezések. Így van ez a hírlapok esetében is, bár nagy tömegű előfordulásuk nem jellemző az EPA-ban, a MEK időszaki kiadványokra szakosodott változatában.

Ha a hírlapokra mint dokumentumtípusra keresünk, és beállítjuk az archivált kiadványokat (ezeknél a fájlok is megtalálhatóak az EPA szerverén), akkor 35 találatot kapunk, míg magára a hírlap kategóriára 215 találatot, amelyben benne vannak az archivált, a távoli és az offline hozzáférési kiadványok is (ezt a három hozzáférési szintet használja az EPA). Eltöprenghetünk azon, hogy miért éppen ezek a hírlapok kerültek az EPA-ba, ki és mikor finanszírozta a digitalizálásukat, és egyáltalán miért éppen ezeknek az újságoknak a digitalizálása és szolgáltatása merült fel?

A találatokból most csak a legérdekesebbekkel foglalkozunk, amelyek egyúttal bemutatnak néhány tipikus bekerülési utat. Az egyik a „Városi lapok” digitalizálási projekt volt 2009-ben, amely az Oktatási és Kulturális Minisztérium állományvédelmi digitalizálási projektje volt városi és megyei könyvtárak számára. Ekkor az OSZK mikrofilm állományából digitalizáltattak pályázati forrásból fontosabb vidéki lapokat. A tervek szerint a résztvevő könyvtárak dolgozták volna fel az anyagot az OSZK segítségével, de a program csak részben valósult meg. A szkennelt anyagok közül sokat eltettek „jobb” időkre, ugyanis nem volt elegendő kapacitás a teljes digitalizált anyag feldolgozására és szolgáltatására. Az OSZK nyilvántartása²⁷ szerint 116 cím került a programba, de bizonyos digitalizált alapanyagok máig nincsenek teljesen feldolgozva. (Ezek egy része nem is alkalmas a feldolgozásra, mert a mikrofilmes forrás miatt minőségi problémák merültek fel.) Ezen az úton került az EPA-ba például néhány váci, pápai, csongrádi, vásárhelyi, kőszegi, gyöngyösi stb. újság. Sajnos, az elkészült kiadványok legnagyobb részét nem az EPA szolgáltatja, hanem a projektben részt vevő helyi szereplők. Ez a megoldás, bár érthető, de problémás, mert így ezek az egyébként rendkívül értékes tartalmak szétaprózódnak, és előfordulhat, hogy az idők folyamán egy

részük nem vehető igénybe (a tovább nem működő adatbázisok vagy egyéb technikai, illetve gazdasági nehézségek miatt).

Egy másik bekerülési kört képvisel az 1956-os napilapok csoportja. Ez a válogatás mennyiségére nem sok (néhány fontos napilap forradalom alatti számairól van szó), ugyanakkor elsődleges kutatási forrásként szolgálhatnak a történészeknek. A következő címek példányairól van szó: *Magyar Ifjúság*, *Magyar Nemzet*, *Magyar Szabadság*, *Néplap*, *Népszabadság*, *Népszava*, *Szabad Ifjúság*, *Szabad Nép*, *Szabad Szó*, *Új Magyarország*. A szkennelést 1956 ötvenéves évfordulójára készítette el magánkezdeményezésben az OSZK munkatársa, *Hum István Balázs*, és átadta a MEK-nek. A MEK integrálta az 1956-os virtuális kiállításba²⁸, illetve címenként felvették az EPA-ba. Az egész kezdeményezés alkalmi akció volt, amely jól mutatja, hogy a tenni akarásnak és lelkesedésnek milyen nagy-szerű eredményei lehetnek.

Az EPA gyűjteményéből *Az Est* című bulvárlap 1926. április havi számai személyes kedvenceim lettek. A lapot 2008 végén digitalizálták az OSZK-ban, valószínűleg külső megrendelésre, de pontos információt nem tudtak adni az EPA munkatársai. Remek dolgokat lehet találni ezekben a lapszámokban, és az írások jól rávilágítanak arra is, hogy az élet régen sem volt kevésbé botrányos és könnyebb sem, mint manapság.

• Arcanum Adatbázis Kft.

Ha digitalizálásról van szó, nem hagyható ki az Arcanum Adatbázis Kft. tevékenysége, hiszen mára ez a cég az egyik legjelentősebb szereplővé vált ezen a piacon. *Biszak Sándor* cégvezető szerint havi 300–400 ezer oldalt szkennelnek be. Lehet azon vitatkozni, hogy jó-e az, ha a közgyűjtemények helyett egy magáncég végez ilyen típusú tevékenységet, de az nem tagadható, hogy a teljesítményük elismerésre méltó: olyan kiadványokat digitalizáltak, amelyek még jó ideig nem kerültek volna sorra a közgyűjteményi digitalizáló műhelyekben. Az okokról külön tanulmányt lehetne írni, de az belátható, hogy az Arcanum hiánypótló tevékenységet végez,

értékmentő (állományvédelem) és értékteremtő (hozzáférhetővé tétel) digitalizálásával.

A nagytömegű digitalizálási projektek országonként eltérő minta szerint alakultak, fogjuk fel előnyként, hogy nálunk az Arcanum Kft. által képviselt piaci modell is megjelent a közgyűjteményi digitalizálások mellett: általuk jóval nagyobb tempóban tud gyarapodni a magyar nyelvű és magyar vonatkozású digitalizált tartalom az interneten, mint a szerényebb anyagi erőforrásokkal rendelkező állami intézmények digitalizálási tevékenysége által.

Az utóbbi időben két olyan szolgáltatással jelentkezett a cég, amelynek sikerült elérnie a közgondolkodás ingerküszöbét. Az egyik a Mapipe (<http://mapire.eu/hu>), a másik pedig az Arcanum Postcards (<http://postcards.arcanum.hu/hu>). Mindkettőről lelkesen számoltak be a nagy internetes portálok, és számtalan Facebook megosztást generáltak, a visszajelzések szerint az emberek szeretik és használják ezeket a szolgáltatásokat.

A vállalkozás könyvtárak és egyéni kutatók számára is fontos adatbázisa az *Arcanum Digitális Tudománytár*, amelyben a cég által digitalizált periodikákat szolgáltatják teljes szövegű keresettséggel. A hozzáférés korlátozott, de a nagyobb hazai könyvtárak az EISZ-en keresztül elérhetik.

Az Arcanum Adatbázis Kft. hírlapok digitalizálásával is foglalkozik. Egy 2013-ban tartott konferencián (*Tartalmak a digitális térben – digitális jelen-lét*, 2013. március 19.) Biszak Sándor a Népszava, a Pesti Hírlap, a Pesti Napló és a Budapesti Hírlap digitalizálásáról számolt be.²⁹ A digitalizálandó nyersanyagot könyvtárraktól (többek között az OSZK-tól, az SZTE Klebelsberg Könyvtártól és más együttműködő partnerektől) szerzik be. A hiánypótlás sok esetben nem egyszerű, gyakran több tucat könyvtárat kell végigkérdezni, amíg sikerül pótolni a hiányokat. A hivatkozott előadásból kiderülnek a vállalkozás munkamódszerei is. Amennyiben mód van rá, a bekötött köteteket lapokra vágják, hogy nagysebességű lapadagolás szkennelvel lehessen elvégezni a munkát, ami nagyságrendekkel hatékonyabb (és olcsóbb), mint a ha-

gyományos lapozgató módszer. A felvágott, és könyvtárraktól kölcsönként köteteket digitalizálás után természetesen visszakötik. A szkennelést 300 DPI-ben végzik, amelynek során JPEG formátumú képek keletkeznek, majd az automatikus szövegfelismerés után előállnak a kétrétegű PDF állományok, amelyek idővel bekerülnek a szolgáltatásba.

2014. június végétől elkezdték szolgáltatni a 20. század első felének legjelentősebb magyar hírlapjait az Arcanum Digitális Tudománytáron (ADT <http://adtplus.arcanum.hu/hu/>) keresztül, amelyekből az Index.hu naponta szemléz a már említett Nagy háború rovatában.³⁰ Egy másik, kisebb gyűjteményük a Helyi Lapok névre hallgat, és a <http://www2.arcanum.hu/helyilapok> címen érhető el. A bemutatkozó szöveg szerint „*A Helyi Lapok projekt immár összesen 435 kötetet dolgoz fel a különböző, helyi érdekeltégű, elsősorban napilapok közül. Összesen 181 561 oldal érhető el kereshető formában.*” Elkészült a felvidéki Új Szó 1948 és 1999 között megjelent lapszámainak digitalizálása is, amely a <http://adatbank.sk/digiteka/digitalizalt-folyoiratok/uj-szo> címen érhető el.

• Helyi és egyéni kezdeményezések

Nem mehetünk el a helyi kezdeményezések mellett sem, melyek mögött hol könyvtárak, hol magánszemélyek állnak. Itt megint csak nem tudjuk körbejárni az országos helyzetet, ezért csak néhány kiragadott példát ismertetek. Szerencsére ennél jóval több kezdeményezés létezik országszerte, amelyek mind nagyon fontosak a helyi közösségek helytörténetét illetően. Ez hol egy-egy újságkivágat-gyűjteményt jelent, mint például a nyiracsádi Ligetalja Könyvtár esetében³¹, hol egy-egy JaDoX-os gyűjteményt, amint a szentesi példa³² (*Szentesi Lap 1872–1947*, *Szentesi Élet 1968–*) mutatja. Említhetjük még a nagykanizsai Halis István Városi Könyvtár szolgáltatását³³ is, amelyben korabeli helyi hírlapokat digitalizáltak és szolgáltatnak.

Nem szeretném ilyen szempontból körbejárni az országot, de mindenesetre megnyugtató, hogy sok helyen fontosnak tartják hasonló gyűjtemé-

nyek kialakítását. Ehelyett inkább két témába vágó magánkezdeményezésű blogról számolnék be, amely műfaj véleményem szerint sokat tud tenni a régi digitalizált hírlapok népszerűsítéséért és köztudatban való terjesztésükért. Az egyik egy főként tipográfiával foglalkozó blog, a *Font.hu*. Itt elsősorban a „Tegnap írás” rovat³⁴ érdekes témánk szempontjából, ugyanis különböző napilapok teljes lapszámaikat tartalmazza PDF verzióban. A blogger nem számolt be a motivációjáról, mindenesetre pozitív a kezdeményezés.

Egy másik blog, a *Hasznos mulatságok*³⁵ szerzője viszont leírta motivációját: az oldalt bevalloottan az Osztrák Nemzeti Könyvtár szolgáltatása, a korábban már bemutatott ANNO ihlette, mottójául a következő idézetet választotta: „[a hírek] esetében soha nem az a fontos, hogy megfeleljen-e a valóságnak, hanem csak az, hogy az emberek elhiszik-e azokat. Ha elhiszik, közvéleményt formáló tényezővé válnak.” (Borsányi György). A blog képeket és lapkivágatokat közöl a korabeli újságokból. A bejegyzések címkézettek, ami megkönnyíti a navigációt.

A Délmagyarország digitalizálásának tapasztalatai

A továbbiakban saját munkánkról, a Délmagyarország című napilap digitalizálásáról számolok be. Az egyes munkafolyamatok ismertetésétől eltekintek, mert ennél fontosabbnak gondolom, hogy átfogó képet adjak a projektről. Annál is inkább, mivel a Klebelsberg Könyvtár Youtube csatornáján (<http://www.youtube.com/sztebibl>) elérhető egy szakmai közönségnek szánt videó (Hogyan készült a DélmagyarArchív?), amely lépésről-lépésre bemutatja a teljes munkafolyamatot.

Kezdetek

Több mint egy évszázad történelmét rejti az 1910-ben indult Délmagyarország, mely a vidéki Magyarország egyik legpatinásabb napilapja volt. Napjainkban nem csupán értékes történeti forrás, hanem érdekes olvasmány is, hiszen a

fontos történelmi eseményekről szóló beszámolók mellett a hétköznapok történései is megjelennek az oldalakon. Nagy élmény olvasgatni például az 1920-as évek apróhirdetéseit, de legalább ennyire érdekesek a kor politikai csatározásairól készült írások. A napilap digitalizálásával és kereshetővé tételével, illetve egyes kiemelt cikkek cím szerinti feldolgozásával olyan lehetőségek nyílnak meg az érdeklődő olvasók előtt a régió és Szeged történetének kutatásában, amelyek eddig elképzelhetetlenek voltak.

A digitalizáláshoz szükséges példányokat az SZTE Klebelsberg Könyvtár és a Somogyi-könyvtár közösen adja össze, és a digitalizálás folyamatában is szorosan együttműködnek. Erre az együttműködésre a digitalizálandó anyag mennyisége miatt nagy szükség van, nem beszélve arról, hogy a feldolgozott állomány minél teljesebbé tétele érdekében is elkerülhetetlen az együttműködés.

Az első hírek az újság digitalizálásáról már 2011-ben megjelentek, de a tényleges, nagyléptékű digitalizálási munka csak 2012 vége felé kezdődhetett el, az előzetesen elvégzett tesztek, illetve próbaszkennelések tapasztalatainak összegzése után. Ezt az egy évet nem töltöttük tétlenül, ekkor zajlottak azok az előkészítő lépések, amelyek a későbbi munkálatok gördülékenységét biztosították. Oldalról-oldalra felmértük mindkét könyvtár Délmagyarország állományát, felderítettük az elérhető duplumpéldányokat. Már a projekt kezdetén célként merült fel az eddig analóg formában létező, az újságban található egyes érdekesebb cikkeket feltáró cédulakatalógusok elektronikus formába konvertálása, ezért ezt is meg kellett tervezni.³⁶



A Somogyi-könyvtár helytörténeti cédulakatalógusai, amelyeket digitalizáltunk

tár is. Az elkészült adatbázis kialakításában és szolgáltatásában szorosan együttműködünk a Délmagyarország jelenlegi kiadójával, ők abban tudnak sokat segíteni, hogy a *DélmagyArchív* névre keresztelt archívum minél több érdeklődőhöz eljusson.

Mennyiség

A lap teljes élettörténete 105 évet (1910–2014) fog át. Előzetes kalkulációink szerint, ha egy évben minden nap megjelent a napilap (nyilvánvalóan nem jelent meg), akkor a 105 évfolyam alatt összesen maximum 38 325 szám jelenhetett meg. Valójában átlagosan csak 300 körüli lapszám jelent meg évente, ez is több mint 30 000 lapszámot jelent.

Résztevők

Mivel járultak hozzá a résztvevők a projekt sikeréhez?	
Somogyi-könyvtár	SZTE Klebelsberg Könyvtár
Eszközzel	Eszközzel
Nyersanyaggal	Nyersanyaggal
humán erőforrással	humán erőforrással
helytörténeti adatbázissal (cédulakatalógus)	egyetemtörténeti adatbázissal (cédulakatalógus)
	szoftverfejlesztéssel
	a projekt koordinálásával

Elvégzendő munkák az egyes könyvtárakban	
Somogyi-könyvtár	SZTE Klebelsberg Könyvtár
digitális másolat készítése a lapszámokról	digitális másolat készítése a lapszámokról és a cédulákról
Hiánypótlás	cédulakatalógusok átírása, konvertálása
	OCR, lapszámok ellenőrzése, körbevágása
	digitalizált állományok utómunkái, hiánypótlás
	rekordok generálása az előzetesen elkészített segédfájlok alapján
	konvertált rekordok javítása, újak készítése

A hiánypótlásban segítséget kaptunk az Országos Széchényi Könyvtártól és duplum példányokkal hozzájárult a sikerhez az Országgyűlési Könyv-

Ha egy számot 12 oldalasnak veszünk (ami szintén erős felülbecslés, mivel még 1994-ben is csak 12-16 oldalas számok vannak, 2003-ban

24 oldalast találtunk, az 1920-as évekből 4-6 oldalt, 1955-ből 8 oldalt, 1972-ből 8 oldalt), akkor összesen 459 900 oldalt mutat a terjedelmi becslés ($38\,325 \times 12 = 459\,900$). Ennél pontosabb becslést nem tudunk előzetesen adni, a valós mennyiséget 300–500 ezer oldal közé tettük. Az eltelt időszak tapasztalatai alapján elmondhatjuk, hogy a végső terjedelem inkább az alsó becsült értékhez lesz közelebb. A raktári polcon ez a 105 év körülbelül 20 méternyi bekötött kötetet jelent.

Felmerülő hibatípusok és más problémák

A digitalizálás során számtalan hibatípussal találkoztunk, amiket érdemes összegyűjteni avégett, hogy tapasztalatul szolgáljanak mások számára is, hogy mire érdemes odafigyelni az ilyen típusú digitalizálási projekteknél. A leggyakoribb problémák a következők voltak:

- szakadások,
- kivágott cikkek,
- kitépelt részek,
- teljesen hiányzó oldalak és lapszámok,
- fénymásolattal pótoltt részek,
- elmosódott nyomtatás és más nyomtatási hibák,
- firkálások, foltos, tintás, koszos papír,
- ragasztott papír,
- gyűrött lapok,
- lapszámozási problémák.

A tökéletességre törekedve, szeretnénk volna, ha a közreműködő könyvtárak papírállományánál teljesebb és jobb lesz a végeredmény. Elmondhatjuk, hogy a helyi együttműködésnek és a külső forrásból (pl. OSZK, OGYK) érkező hiánypótlásoknak köszönhetően ez többé-kevésbé sikerült.

Néha komoly bonyodalmat okoztak az elszámozások, ugyanis előfordultak olyan időszakok, hogy több hónapon keresztül rossz lapszámmal jelent meg a napilap, aztán egyszer csak észbe kapott valaki a szerkesztőségben, és akkor visszatértek a helyes lapszámozáshoz. A nyilvánvaló és sűrűn előforduló tévedések miatt nem ragaszkodtunk a címlapokra nyomtatott

lapszámozáshoz (nem is tehetjük volna, ez sokszor komoly keveredést idézett volna elő), hanem próbáltuk rekonstruálni a helyes és időben folyamatos számozást, amelyhez a következő fájlnévsémát használtuk: dm_19xx_00x.pdf. Tehát a lapszámok kerültek bele a fájlnevbe és nem a dátum. Természetesen először felmerült, hogy az egyes lapszámok megjelenési dátumát kellene használni a fájlnevben (hiszen ez egyedi és ritkábban fordult elő körülötte az elnyomtatásból származó tévedés), azonban ezt elvetettük, mert így nehezebb lett volna a hiányzó lapszámok kiszűrése, hiszen az újság nem feltétlenül jelent meg minden nap. Nagy általánosságban csak hétfőnként nem jelent meg, azonban ahhoz sűrűn fordultak elő más napokon nem megjelenő lapszámok, hogy e fölött szemet hunyjunk. Ezzel szemben a lapszámozásnak elméletileg folyamatosnak kell lennie, így ebben az esetben könnyebben kiszűrhető, ha egy-egy sorszámugrás fordul elő.

Eredmények

Az eddig feldolgozott anyag a számok tükrében a következőképpen alakult:

Feldolgozott évfolyamok: 1910–1945 (36 évfolyam)

Lapszámok: 10 344 db

Oldalszám: 99 079 oldal

Analitikusan feltárt cikkek száma: 17 344 rekord az egyetemtörténeti cikkadatbázisból + 19 189 rekord a helytörténeti cikkadatbázisból + 24 580 db katalóguscédula, összesen 61 113 tétel. Ebből az 1910–1945 időszakra vonatkozó, tehát jelenleg már használható rekordok száma: 19 854. Duplumszűrt, kulcsszó-alakú kifejezések száma a metaadatokban: 2 960 988 és az OCR szövegben: 72 701195 db.

1910	1	6	12	1911	1	6	12	1912	1	6	12	1913	1	6	12	1914	1	6	12
1915	1	6	12	1916	1	6	12	1917	1	6	12	1918	1	6	12	1919	1	6	12
1920	1	6	12	1921	1	6	12	1922	1	6	12	1923	1	6	12	1924	1	6	12
1925	1	6	12	1926	1	6	12	1927	1	6	12	1928	1	6	12	1929	1	6	12
1930	1	6	12	1931	1	6	12	1932	1	6	12	1933	1	6	12	1934	1	6	12
1935	1	6	12	1936	1	6	12	1937	1	6	12	1938	1	6	12	1939	1	6	12
1940	1	6	12	1941	1	6	12	1942	1	6	12	1943	1	6	12	1944	1	6	12
1945	1	6	12	1946	1	6	12	1947	1	6	12	1948	1	6	12	1949	1	6	12
1950	1	6	12	1951	1	6	12	1952	1	6	12	1953	1	6	12	1954	1	6	12
1955	1	6	12	1956	1	6	12	1957	1	6	12	1958	1	6	12	1959	1	6	12
1960	1	6	12	1961	1	6	12	1962	1	6	12	1963	1	6	12	1964	1	6	12
1965	1	6	12	1966	1	6	12	1967	1	6	12	1968	1	6	12	1969	1	6	12
1970	1	6	12	1971	1	6	12	1972	1	6	12	1973	1	6	12	1974	1	6	12
1975	1	6	12	1976	1	6	12	1977	1	6	12	1978	1	6	12	1979	1	6	12
1980	1	6	12	1981	1	6	12	1982	1	6	12	1983	1	6	12	1984	1	6	12
1985	1	6	12	1986	1	6	12	1987	1	6	12	1988	1	6	12	1989	1	6	12
1990	1	6	12	1991	1	6	12	1992	1	6	12	1993	1	6	12	1994	1	6	12
1995	1	6	12	1996	1	6	12	1997	1	6	12	1998	1	6	12	1999	1	6	12
2000	1	6	12	2001	1	6	12	2002	1	6	12	2003	1	6	12	2004	1	6	12
2005	1	6	12	2006	1	6	12	2007	1	6	12	2008	1	6	12	2009	1	6	12
2010	1	6	12	2011	1	6	12	2012	1	6	12	2013	1	6	12	2014	1	6	12

Contentában Szakaszhatár
 SCAN kész NEM jelent meg
 Vágás/pótlás alatt
 Betöltésre kész
 KK hiány

Az előrehaladás nyomon követésére kitalált rendszer

Mit tartalmaz az adatbázis?

Jelenleg a repozitórium az 1910 és 1945 között megjelent Délmagyarország-lapszámok digitális verzióját tartalmazza kétrétegű PDF formátumban, ezzel lehetővé téve a fájlokban való keresést, ugyanakkor visszaadva az eredeti lapképet is.

Az adatbázisban a teljes szövegű PDF-ek mellett részletesen feldolgozott cikkekordokat is hozzáférhetővé tettünk. Ezek több forrásból kerültek és kerülnek be a jövőben is a rendszerbe. Egy részük eddig katalóguscédulákon volt hozzáférhető (pl. a Somogyi-könyvtár helyismereti katalógusának Délmagyarországra vonatkozó céduláin), ezért ezeket előbb át kellett írni elektronikus formába. Más részük több elektronikusan létező adatbázis összefésülésével került be, ilyen források voltak például az előző pontban említett egyetemtörténeti cikkadatbázis vagy a helytörténeti cikkadatbázis.

Így a több szempontú válogatás (helytörténet, egyetemtörténet, személynevek, tárgyszavak) miatt duplumok is előfordulhatnak. Fontos kiemelnünk, hogy a rekordok bővítése, javítása folyamatosan zajlik, hiszen amikor minden be-

töltésre kerül, akkor nagyjából 150 ezres rekord-számról beszélhetünk majd.

Böngészés

© Lapszámok szerint

Az adatbázisban természetesen lehetőség van az időben egymást követő lapszámok böngészésére a nyomtatott számoknál megszokott módon. Ezt a módszert akkor érdemes választani, ha nem konkrétan keresünk valamit, hanem csak megszeretnénk nézni például az 1912-es év lapszámaikat. Ezt a nézetet érdemes abban az esetben használni, amikor egy konkrét esemény előzményeit/magát az eseményt/az esemény következményeit szeretnénk vizsgálni. (Például: kíváncsi vagyok az 1928-as év apróhirdetéseire; kíváncsi vagyok hogyan számolt be a Délmagyarország a második világháború kitöréséről; a nagypám születésnapján megjelent számot szeretném megnézni stb.)

© Cikkek szerint

Ez a böngészési módszer alkalmas leginkább arra, hogy átfogó képet kapjunk arról, mi található az egyes évek lapszámaiban. Szintén ez a típusú böngészés alkalmas arra, hogy érdekessé-

gekre akadunk a feldolgozott rekordok között. (Például: szeretnék átfogó képet kapni arról, mi történt 1930-ban; szeretnék valamilyen érdekességet/meghökkenítő/vicces dolgot találni a régmúltból stb.)

Keresés a dokumentumok adataiban

🕒 Cím

Ez az egyik legcélravezetőbb keresési megoldás, ha valamilyen konkrét dologra vagyunk kíváncsiak. Eredményül azokat a találatokat kapjuk, amelyek tartalmazzák azt a kifejezést, amire rákerestünk. (Például: a fogadalmi templom felépítésével kapcsolatos cikkeket keressük; *Somogyi Szilveszter*ről szóló cikkeket keressük stb.)

👤 Szerző

Akkor alkalmazzuk, ha egy szerző cikkére, cikkeire vagyunk kíváncsiak. Tudnunk kell azonban, hogy a napilapoknál bevett gyakorlat volt, hogy nem mindig szerepeltetik a cikk íróját, van hogy csak monogramot használnak, ezekben az esetekben természetesen az adatbázisban sem visszakereshetőek a szerzőségi adatok. (Például: a *Péter László* által írt cikkeket szeretnénk kilistáztatni.)

🗄 Tárgyszó

A címre való keresés mellett, ha konkrét témára keressük, tárgyszót érdemes használni. Akkor vezet célra, ha a keresési kulcsszó valami miatt nem szerepel a címben, ugyanakkor a cikk mégis az adott témáról szól. Például: *Szent-Györgyi* kulcsszóra a címben keresve 345 találatot kapunk (ennyi cikk címében szerepel a kulcsszó), míg ugyanerre a kulcsszóra keresve a tárgyszó mezőben keresve 339 találatot kapunk (ennyi rekordhoz van felvéve a *Szent-Györgyi* tárgyszó). A két találati lista láthatóan nem egyezik meg!

📅 Dátum

Egy adott napon, vagy egy adott évben megjelent cikkek kérhetőek le ezzel a funkcióval, hasonlóan a 'Böngészés' részben felvázolthoz. Természetesen az egyes keresési ismérvek kombinálhatóak egymással, így válik az lehetővé, hogy az 1923-as év *Juhász Gyulával* kapcsola-

tos cikkeit könnyen megtalálhassuk (róla szól, ő írta, a neve szerepel a címben). A megjelenés éve intervallumosan is kereshető (1936–1938). (Például: egy adott napon megjelent újságot szeretnénk kikeresni; az 1923-as év *Juhász Gyulával* kapcsolatos cikkeket keressük; olyan cikkeket keressük, amelyeknek szerepel a címében 'Móra' neve és 1920 és 1930 között jelentek meg.)

📄 Teljes szöveg

Nem összekeverendő a *Keresés a dokumentumok teljes szövegében* funkcióval! Ha a *Keresés a dokumentumok adataiban* van kiválasztva és a lenyíló menüben választjuk ki a *Teljes szöveg* sort, akkor a dokumentumok adatainak a teljes szövegében fogunk keresni (szerző, cím, tárgy-szó, megjelenési adatok, megjegyzések, stb.), és nem a digitalizált állományok teljes szövegében! Így válnak egyben kereshetővé a feldolgozott cikkek rekordok az imént zárójelben felsorolt adatai. (Például: így kapjuk a legteljesebb találati halmazt például a *Szent-Györgyi* kulcsszóra is, 508 találattal, amely minden olyan rekordot tartalmaz, ahol ez a kulcsszó a feldolgozott rekordokban valahol előfordul.)

Keresés a dokumentumok teljes szövegében

A szkennelt oldalak felismerésére automatikus karakterfelismerést alkalmaztunk, így előfordulhatnak rosszul felismert karaktersorozatok, azonban a korai évek kivételével (ezeknél az OCR alacsonyabb hatásfokú a rosszabb nyomdakép miatt) elmondható, hogy az OCR pontossága nagyjából eléri a 90–95%-ot.

Például: Szeged egyik városrészével, Felsővárossal kapcsolatos anyagokat keressük. Amennyiben a dokumentumok adataiban való keresés is bekapcsolva marad, akkor ezek a találatok vannak a találati lista elején, míg a teljes szövegű találatokkal folytatódik az eredménylista. A teljes szövegű keresés alapvető tulajdonságából fakadóan, ha valamilyen általános dologra keressük, akkor nagyon sok találatot kapunk. Ezt bizonyos esetekben érdemes finomítani, de ha erre nincsen mód, akkor marad a találatok végignézése. Példánk esetében sem nagyon van mód

a 'Felsőváros' keresőszó módosítására, esetleg időbeli szűkítéssel tudunk próbálkozni.

Egy másik példa: *Szőkefalvi-Nagy Gyula* matematika professzorral kapcsolatos találatok érdekelnek bennünket. A többtagú keresőszót idézőjelben kell beírni a keresőmezőbe („*Szőkefalvi-Nagy Gyula*”), amennyiben azok a találatok érdekelnek bennünket, amelyekben egymás után fordulnak elő a kifejezés, vagy jelen esetben a név szavai. Ellenkező esetben megkapunk minden olyan találatot, ahol a kifejezés tagjai egy oldalon belül előfordulnak. Természetesen ez is hasznos lehet bizonyos esetekben, de példánk esetében az idézőjeles keresésre van szükségünk, mivel specifikus és direkt információt szeretnénk a fent nevezett személyről.

Web2.0-ás megoldások

A projekt során hamar felmerült az igény, hogy a felhasználókat minél inkább megpróbáljuk majd bevonni a Délmagyarország feldolgozása körüli munkálatokba. Úgy gondoljuk, hogy volna igény ilyen típusú szolgáltatásra, egyúttal a tömegek bölcsessége sok pluszt tudna hozzatenni a digitalizált oldalakhoz. Az ilyen típusú „crowdsourcing”-nak azonban komoly szoftveres feltételei vannak, így a 2014. június közepén elindult DélmagyArchív szolgáltatás 1.0-ás verziójában még nem szerepelnek ezek a lehetőségek.

Szükséges azonban annak a megtervezése és felmérése, hogy egy későbbi fejlesztés során egyáltalán milyen Web2.0-ásnak mondott szolgáltatásokat volna érdemes biztosítanunk az archívum használói számára. Ez a fejezet tehát mintegy vázlatként, gondolatkísérletként ezt hivott össze gyűjteni.

👁 Cikk-rekordok létrehozása

Annak a lehetőségnek a biztosítása, hogy a már a rendszerben található analitikusan feldolgozott rekordokhoz hasonlókat létrehozassanak a regisztrált felhasználók. Ez a legkomolyabb szint, így fontos, hogy az általuk létrehozott rekordok először egy munkaterületre kerüljenek, ahol egy hozzáértő feldolgozó könyvtáros el tudja bírálni, bekerülhet-e az éles archívumba az adott leírás.

Ha vannak, akkor azonnal ki tudja javítani a kisebb hibákat, ellenőrizni tudja a rekord és digitalizált állomány kapcsolatát.

👁 Cikk-kivágatok készítése

Komoly fejlesztést igényelne egy ilyen funkció bevezetése, hiszen jelenleg nem a cikk a legkisebb egység a digitalizált tartalmat tekintve, hanem a teljes oldal. Több bemutatott adatbázisban alkalmazzák azonban a cikkekre bontást, többnyire a METS/ALTO szabvány segítségével. A „kivágott” cikkeket saját gyűjteménybe lehet rendezni, elküldeni, kinyomtatni, megosztani, címkézni, stb. lehet őket.

👁 Kommentelési lehetőség

Elképzelhető, hogy egyes analitikus rekordok, oldalalak, kigyűjtött tematikus válogatások (címkék), kivágatok, fotók akkora érdeklődést váltanak ki, hogy a használók hozzá szeretnének szólni az adott elemhez. Például egy korabeli május elsejei felvonulásról beszámoló cikkhez plusz információt tudnának hozzatenni az azon résztvevő emberek, ezt tudná biztosítani ez a funkció, azonban itt is szükséges az előzetes moderáció megteremtése, hogy ne lehessen teleszemetelni az adatbázist, mert a mai magyar internetvilóság azt bizonyítja, hogy szükség van az előzetes kontrollra. Ezt a feladatot el tudná látni az archívumot gondozó feldolgozó vagy a tájékoztató könyvtáros.

👁 OCR hibák javítása

Az automatikus karakterfelismerés minőségét nagyban befolyásolja az alapanyag minősége. Ez sokszor már a nyomtatott számban sem volt tökéletes, és a hiba a szkennelés során reprodukálódott. Becsléseink alapján az OCR hatásfoka a teljes adatbázisra vetítve eléri azért a 95%-ot. Ezt tovább lehetne javítani olyan funkció biztosításával, ahol az adatbázis használói (legyenek akár elkötelezett felhasználók, vagy gyakorló könyvtárosok, könyvtáros hallgatók) egy felhasználóbarát felületen javíthatnák a rendszerben az egyes lapképekhez tartozó indexállományt. A nemzetközi példánál bemutatott ausztrál rendszer tudja ezt a funkciót, és láthatóan nagy sikernek örvend használói körében.

👁 Címkézés

A tartalmi feltárást segítő funkció. Amit a weben címkének hívnak, sok esetben a könyvtárosok tárgyszónak nevezik, de nagyjából ugyanarról a dologról van szó. Amennyiben biztosítanánk a felhasználóink számára ezt a lehetőséget, akkor idővel akár komoly tematikus válogatások is készülhetnének az egyes különálló felhasználók címkéinek csoportosításával. Például ha tíz ember létrehoz egy-egy *Klauzál tér* címkét és hozzákapcsol egy-egy cikket, akkor létrejön egy *Klauzál tér* címkénk, amely mögött ott lesz a tíz kigyűjtött cikk.

👁 Tematikus cikkfeldolgozás

Hasonló, mint az előző pont, csak éppen kicsit más oldalról megközelítve. Például, ha valaki rászánja magát, akkor a teljes szövegű keresés segítségével kigyűjtheti, feldolgozhatja és címkézheti az összes Pick szalámigyárral kapcsolatos cikket. Így egyrészt létrejöhet egy *Pick* címke, amely mögött ott lesznek az ezzel a címkével rendelkező rekordok, illetve azt is látni kell, hogy ez a típusú funkció jóval szélesebb lehetőségeket biztosít, mint a sima címkézés, hiszen itt teljes kitöltött rekordok jönnek létre, amelyeknek csak egyik eleme a címke (tárgyszó).

👁 Felhasználói aktivitások nyomon követése

Egyrészt a felhasználók motiválása (toplisták), továbbá a moderáció és a párbeszéd lehetőségének megteremtése miatt volna szükség felhasználó-kezelésre. A Web2.0-ás interakciók során bevett gyakorlat, hogy az aktivitások regisztrációhoz vannak kötve, ez tulajdonképpen megkerülhetetlen a fent vázolt okok miatt. Ezért az itt összegyűjtött Web2.0-ás funkciók nagy részéhez előfeltétel volna ezen funkció biztosítása.

Széles körű megosztási lehetőségek beépítése. A szokásos „social”, közösségi funkciók és megosztási lehetőségek egyrészt a szolgáltatás reklámja miatt szükségesek, másrészt pedig azért, mert már annyira beépültek az internetes oldalakba, hogy a felhasználók elvárják jelenlétüket. Amennyiben a Web2-es lehetőségeket maradéktalanul ki akarjuk használni, akkor kívánatos lenne minden adatbáziselem (cikkre-

kord, oldal, címke) mellé odatenni a szokásos megosztási gombokat.

👁 Önálló Facebook oldal létrehozása

Vége egy olyan pont, amely nem igényel további teendőt, hiszen ez már a start idején is működött. Elsősorban a szolgáltatásról szóló hírek összegyűjtésére és Facebookos terjesztésére szolgál, másodsorban pedig önálló tartalmak posztolására (érdekes, vicces cikkek, évfordulók, stb.). Az intézményi Facebook oldalak mellett a külön Facebook oldal létrehozását indokolta többek között a három közreműködő szervezet megléte, illetve a brandképzés (DélmagyarArchív) lehetősége.

👁 Fórum

A felvázolt kommentezés és a Facebook oldal hozzászólási lehetősége mellett nem biztos, hogy szükség van erre a funkcióra, de a Web2.0-ás megoldások teljességének igénye miatt szerepel a felsorolásban. A sokat emlegetett ausztrál példa esetében létezik ez a funkció, és a közreműködő felhasználók és a szolgáltatás üzemeltetői közötti kommunikációra szolgál. Fontos kiemelni, hogy moderációra természetesen ebben az esetben is szükség volna.

A DélmagyarArchív jövője

A Szegedi Tudományegyetem Klebelsberg Könyvtára és a Somogyi-könyvtár jóvoltából 2014. június 12-én elindult Magyarország első, teljes szövegűen kereshető, szabadon hozzáférhető napilaparchívuma, a DélmagyarArchív.³⁷ A szélesebb olvasói kör elérésének szándéka miatt a projekt főbejárata a régiós internetes hírportálként üzemelő Délmagyar.hu alatt található, a www.delmagyar.hu/delmagyarchiv címen, azonban a professzionális felhasználók számára létezik egy másik bejárat is, többek között bővített keresési lehetőségekkel és egy más felülettel, ez a <http://dm.ek.szte.hu> webcímen keresztül érhető el.

A munka az archívum elindulása után sem állt meg, hiszen nagyon sok dolgunk van még. A szkennelési munkálatok 1980-ban járnak, több fronton is folyamatban van a cédulakatalógusok átirása és betöltésük a rendszerbe (illetve a betöltött rekordok ellenőrzése és javítása). A rendszeren is csiszolni szeretnénk majd az első felhasználói tapasztalatok és többek között az előző részben vázolt Web 2.0-ás megoldások alapján.

A júniusi start után a következő szakaszt, amely az 1970-es szegedi nagy árvízig tart, év végén tervezzük átadni (hiszen ekkor lesz tízéves könyvtárunk új épülete, ezt pedig illik megünnepeelni).

Jegyzetek

1. Ősszel indul a Délmagyarország digitalizálása: 102 év, félmillió újságoldal, 2 millió adat – http://www.delmagyar.hu/szeged_hirek/osszel_indul_a_delmagyarorszag_digitalizalasa_102_ev_felmillio_ujzagoldal_2_millio_adat/2297207 [2014. ápr. 9.]
2. SÁNDOR Ákos: A Délmagyarország című napilap digitalizálása Szegeden, Networkshop 2014, Pécs, –http://videotorium.hu/hu/recordings/details/8481,A_Delmagyarorszag_cimu_napilap_digitalizalasa_Szegeden. [2014. okt. 6.]
3. Digitalizálás és elektronikus hozzáférés K2 Kísérleti műhely – <http://katalist-to-feed.blogspot.hu/2014/09/katalist-k2-meghivo.html> [2014. okt. 6.]
4. Digitalizálják a 100 éves Délmagyarországot – http://www.delmagyar.hu/szeged_hirek/digitalizaljak_a_100_eves_delmagyarorszagot/2214199 [2014. ápr. 9.]
5. Magyar Nagylexikon 9. kötet, Magyar Nagylexikon Kiadó, Budapest, 1999, 500. p.
6. Könyvtárosok kézikönyve 1. kötet. Szerk. Horváth Tibor, Papp István. Bp. Osiris Kiadó, Budapest, 1999, 182. p.
7. Index – Tudomány – Történelem – A nagy háború – <http://index.hu/tudomany/tortenelem/anagyhaboru> [2014. okt. 6.]
8. BUZINKAY Géza: Bulvárlapok a pesti utcán = Budapesti negyed. Lap a városról, 5. évf. 1997. 2–3. sz. (16–17.) sz. 31–44. p. – <http://epa.oszk.hu/00000/00003/00014/buzinkay.htm> [2014. okt. 20.]
9. Országos Széchényi Könyvtár Születésnap Újság – http://www.oszk.hu/szulesnap/szines/sz_nyit.html [2014. okt. 6.]
10. Születésnap Újság Somogyi-könyvtár – <http://www.sk-szeged.hu/szulesnap-ujzag> [2014. okt. 6.]
11. Crossoverek (újság) – <https://plus.google.com/photos/116472086794963776491/albums/5370865631009460305> [2014. okt. 06.]
12. The famous newspaper that was used in movies and TV shows for decades - <http://scinotions.com/2013/12/the-famous-newspaper-that-was-used-in-movies-and-tv-shows-for-decades> [2014. okt. 6.]
13. PÉLI Péter: Történelemhamisítás digitalizálással? A digitalizálás rejtett politikája <http://www.nyest.hu/hirek/a-digitalizalas-rejtett-politikaja> [2014. okt. 6.]
14. The unique challenges of newspaper digitization - <http://www.dlconsulting.com/digitization/the-unique-challenges-of-newspaper-digitization> [2014. okt. 6.]
15. Best Practices for Newspaper Digitization - http://www.library.illinois.edu/dcc/bestpractices/chapter_04_newspaperdigitization.html [2014. okt. 6.]
16. The National Digital Newspaper Program (NDNP) Technical Guidelines for Applicants - http://www.loc.gov/ndnp/guidelines/NDNP_201416TechNotes.pdf [2014. okt. 6.]
17. National Digital Newspaper Program Guidelines & Resources - <http://www.loc.gov/ndnp/guidelines> [2014. okt. 6.]
18. KING, Ed: The British Library newspaper collection and the British Newspaper Archive = Refer, vol. 29. 2013. no. 3. 33–36. p.
19. QUAM, Allison: How one library digitized its community's newspapers = American Libraries Magazine, vol. 42. 2011. no. 7–8. 40–41. p.
<http://www.americanlibrariesmagazine.org/article/how-one-library-digitized-its-community%E2%80%99s-newspapers> [2014. okt. 20.]
20. KOKAS Károly: A periodika-digitalizálás és online szolgáltatás mai gondjai és a szegedi SZTE Klebelsberg Könyvtár gyakorlata = Iskolakultúra, 2014. 3. sz. 84–94. p.
21. VEROSZTA Zsuzsanna: Kerekasztal-beszélgetés a digitalizációról, az emlékméntésről és az ezzel együtt járó szelekcióról. Interjúpartnereink: Biszak Sándor, Gárdos Judit, Ládi László, Mikó Zsuzsanna, Nagy Péter Tibor = Educatio, 22. évf. 2013. 3. sz. 405–412. p.
22. http://epa.oszk.hu/01500/01551/00065/pdf/EPA01551_educatio_13_03_2013-3-teljes.pdf [2014. okt. 20.]
LANDGRAF, Greg: Google Discontinues Initiative to Digitize

- Newspapers' Archives = American Libraries, vol. 42. 2011. no. 7–8. 17. p.
<http://www.questia.com/magazine/1G1-264923163/google-discontinues-initiative-to-digitize-newspapers> [2014. okt. 20.]
23. Data Visualization: Journalism's Voyage West
http://www.stanford.edu/group/ruralwest/cgi-bin/drupal/visualizations/us_newspapers [2014. okt. 6.]
24. Purchase a Package and Register <https://www.britishnewspaperarchive.co.uk/account/register?countrykey=0&voucherclaimant=false&gift=false&nextpage=%2fpayments&rememberme=false&cookietracking=false&partnershipkey=0&newsletter=true®isterreason=none> [2014. okt. 6.]
25. Australian Newspaper Digitisation Program – <http://www.nla.gov.au/content/newspaper-digitisation-program> [2014. okt. 6.]
26. Australian Newspaper Plan – <http://www.nla.gov.au/australian-newspaper-plan> [2014. okt. 6.]
27. Indigo - <http://mekosztaly.oszk.hu/indigo/osszetettbev.php?ho1=nev&mit1=&muvelet1=and&hol2=projekt&mit2=V%C3%A1rosi+lapok&muvelet2=and&hol3=nev&mit3=&muvelet3=and&hol4=nev&mit4=&view=list&page=1&tab=all> [2014. okt. 6.]
28. Újságok 1956-ból – <http://mek.oszk.hu/kiallitas/1956/html/ujsgok.htm> [2014. okt. 6.]
29. BISZAK Sándor: Új dimenzió a tartalomszolgáltatásban (Arcanum Digitális Tudománytár 2, napilapok és 3D:s történelmi térképek) – <http://www2.arcanum.hu/hirek/Biszak.ppsx> – [2014. okt. 6.]
30. Olvasson százéves napilapokat ingyen! - http://index.hu/tudomany/tortenelem/anagyhaboru/2014/06/27/indul_az_index_es_az_arcanum_egyuttmukodese [2014. okt. 6.]
31. Újságcikkek - <http://nyiracsadkonyvtar.hu/katalogusok/ujsgcikkek> [2014. okt. 6.]
32. E-könyvtár Szentes digitális helytörténeti gyűjtemény – <http://ekonyvtar.vksz.hu/jadox> [2014. máj. 2.]
33. Folyóiratdigitalizálás – <http://folyoiratok.nagykar.hu> [2014. okt. 6.]
34. Tegnap újság, Font.hu – <http://www.font.hu/b2/?cat=4> [2014. okt. 6.]
35. Hasznos multságok – <http://hasznosmultságok.blogspot.hu> [2014. okt. 6.]
36. KOKAS Károly: A periodika-digitalizálás és online szolgáltatás mai gondjai és a szegedi SZTE Klebelsberg Könyvtár gyakorlata = Iskolakultúra, 2014. 3. sz. 84–94. p.
37. Elkészült a Délmagyarország 1945 előtti lapszámainak digitális archívuma <http://www.hirado.hu/2014/06/12/elkeszulta-delmagyarorszag-1945-elotti-lapszamainak-digitalis-archivuma> - [2014. okt. 6.]

Beérkezett: 2014. október 6.

Könyvtár- és Információtudományi Tanszék Londonban – Mi ebben a meglepetés?

Könyvtár- és Információtudományi Tanszék

(<http://www.city.ac.uk/department-library-information-science>) alakult

a City Egyetemen, Londonban. Elődje az egyetem Lyn Robinson és (főként) David Bawden nevével fémjelezett, nagynevű Információtudományi Tanszéke volt.

Továbbra is kapcsolódik hozzá az Információtudományi Központ (<http://www.city.ac.uk/department-library-information-science/centre-for-information-science>) névre hallgató kutatóközpont.

Blogbejegyzésében (<http://theoccasionalinformationist.com/2014/11/19/on-being-library-and-information-science/>) Bawden arról ír, hogy több mint harminc évvel ezelőtt éppen a többi tanszéktől való különbözőségüket kívánták hangsúlyozni Információtudományi Tanszék elnevezéssel még akkor is, ha hallgatóik jelentős részben könyvtárosként helyezkedtek el. A mostani változtatás, vagyis a könyvtártudomány beemelésének több oka van. Miközben más intézmények az információtudományi jelleget hangsúlyozzák, sőt az iSchool nevet vették fel, a City oktatói megint csak különbözni kívánnak. A fő ok azonban az, hogy hisznek a könyvtárban.

(Koltay Tibor)