

Visky Ákos László

Gyorsmérleg az OSZK Webarchívum és a KDS-K pályázat nyerteseinek együttműködéséről¹

Most, hogy a végéhez értünk a Közgyűteményi Digitalizálási Stratégia könyvtári ágának (KDS-K) keretében megvalósult együttműködésnek – mely során az Országos Széchényi Könyvtár (OSZK) Webarchívuma és a pályázaton nyertes megyei hatókörű városi könyvtárak elsősorban a nemzeti webtér regionális vonatkozású webhelyeinek feltárásán dolgoztak –, illő rövid összefoglalót adni a közös munka eredményéről.

A háttér

„A Közgyűteményi Digitalizálási Stratégia (KDS) célja, hogy biztosított legyen a nemzeti kulturális kincseink, a közgyűteményi tartalmak minél szélesebb körű, akadálytalan hozzáférése a kultúrafogyasztók számára.” – definiálja a Digitális Jólét Program honlapja a projektet, és az abban megfogalmazott célok érdekében feladatokat is meghatároz: „2020-ig a közgyűtemények digitalizált állományának – a teljes digitalizálandó állományhoz képest – meg kell közelítenie az 50%-ot”, „a digitális tartalmak egy közös kereső felületen legyenek elérhetőek”, és „40%-kal emelkedjen a közgyűteményi tartalmakra épülő digitális tananyagok száma”.² Azonban a stratégia nemcsak az analóg dokumentumok digitalizálásáról és hozzáférhetőségéről szól, hanem újfajta tevékenységként megjelenik benne az interneten elérhető tartalmak megőrzésének az igénye, a webarchiválás is.

A Közgyűteményi Digitalizálási Stratégia (2017–2025) című dokumentum³ a magyar digitális örökség megőrzésének eredményeinek elemzésénél még mint egy rendezetlen területtel foglalkozott a webarchiválással, de együttal kiemelte az on-

line megjelenő források megőrzésének, visszakereshető webarchívumokban való tárolásának és bárki számára hozzáférhetővé tételének fontosságát;⁴ és célként jelölte meg a KDS első ütemében (2017–2020) az aktív webaratás jogi és technikai feltételeinek megteremtését, valamint a második ütemében (2021–2025) az aktív webaratás fokozatos kiterjesztését⁵ és szolgáltatásként történő működését Kárpát-medencei hatókörrel.⁶

Bár a KDS-projekt könyvtári ágazatának aggregátora az OSZK, a könyvtári dokumentumok digitalizálásának összefogása mellett a webarchiválási tevékenységével is érintett: a webarchiválás fontosságának felvetését⁷ elfogadva, azt már korábban fejlesztési céljai közé emelte a nemzeti könyvtár – és bekerült a 2015-ös gyűjtőköri szabályzatába⁸ is –, azonban a megvalósulásra csak a 2017-ben indult Országos Könyvtári Rendszer Projekt⁹ (OKR-projekt) keretében nyílt lehetőség. Azóta az OSZK-ban folyamatosan teremődik meg a feltételrendszere a webarchiválásnak, a technikai eszközök beszerzése mellett létrejött egy kísérleti munkacsoport¹⁰ is, mely az elmúlt években szisztematikusan építkezve kialakította egy üzemszerűen működő webarchívum alapjait.¹¹ 2020 májusában az Országgyűlés a kulturális törvény módosításával¹² a nemzeti könyvtár feladatává tette a webtartalom megőrzését – a részleteket rendelet fogja szabályozni –, megteremtve ezzel a tevékenység jogi háttérét, júniusban pedig életbe lépett az OSZK új szervezeti és működési szabályzata,¹³ mely önálló osztályt hozott létre a webarchiválás feladatainak ellátására.

A KDS végrehajtásának keretében 2019-ben a könyvtárak együttműködését segítendő, a megyei hatókörű városi könyvtárak (Pest megye és Budapest területén lévő könyvtárak kivételével), állami egyetemi könyvtárak, illetve azok konzorciumai számára írtak ki pályázatot a könyvtári digitalizálás támogatására, kiemelten helytörténeti vonatkozással.¹⁴ A digitalizálás mellett a pályázók feladata volt „az OSZK webarchiválási tevékenységének segítése, elsősorban a helyi vonatkozású oktatási, tudományos és kulturális online tartalmak esetében. Részfeladatok: archiválendő webhelyek válogatása, jogtulajdonosok elérhetőségének kiderítése, a mentett weboldalak minőségének ellenőrzése és metaadatok készítése.”¹⁵

A KDS-K pályázat keretében az OSZK webarchiváló munkacsoportja az alábbi célokat tűzte maga elé:¹⁶

- hosszú távú együttműködés megalapozása a pályázatban nyertes könyvtárakkal a webarchívum gyarapításában és egyéb munkafázisokban;
- írásos anyagok és előadások formájában módszertani segítség nyújtása a partnerintézményeknek az online tartalmak archiválásával kapcsolatosan;
- szakmai ismeretterjesztéshez a közoktatásban is felhasználható mintaalkalmazás készítése, amellyel illusztrálható, hogy a webarchívum anyaga hogyan tudja kiegészíteni a könyvtárak „hagyományos” digitális gyűjteményeit;¹⁷

- a középiskolás korosztálynak szánt elektronikus tananyag összeállítása az intézményi és a személyes webarchiválásról, a digitális kultúránk megőrzésének fontosságáról és lehetséges módszereiről.

2019 októberében született döntés a nyertesekről,¹⁸ de technikai és adminisztratív nehézségek miatt a tényleges munka csak 2020-ban indulhatott meg, és végül az alábbi 11 megyei hatókörű városi könyvtár vehetett részt a projektben:

- Bács-Kiskun Megyei Katona József Könyvtár, Kecskemét
- Balassi Bálint Megyei Könyvtár, Salgótarján
- Békés Megyei Könyvtár, Békéscsaba
- Berzsenyi Dániel Megyei Hatókörű Városi Könyvtár, Szombathely
- Bródy Sándor Megyei és Városi Könyvtár, Eger
- József Attila Megyei és Városi Könyvtár, Tatabánya
- Méliusz Juhász Péter Könyvtár, Debrecen
- Móricz Zsigmond Megyei és Városi Könyvtár, Nyíregyháza
- II. Rákóczi Ferenc Megyei és Városi Könyvtár, Miskolc
- Somogyi Károly Városi és Megyei Könyvtár, Szeged
- Vörösmarty Mihály Könyvtár, Székesfehérvár

Persze joggal merül fel a kérdés, miért fontos ez az együttműködés, ha a nemzeti könyvtárban már dolgoznak a magyar vonatkozású webtartalom archiválásán, mit tud ehhez hozzátenni egy más erőforrásokkal rendelkező könyvtár? Miért jó ez egy megyei hatókörű intézménynek, vagy miért ne csak egy pályázat kötelezően letudandó köreként tekintsen rá? Pont azért, amiért szükség van helyi könyvtárakra vagy más közintézményekre: helyi beágyazottsága előnyt jelent a különböző munkafolyamatokban, például a térség webhelyeinek feltárásában vagy a kapcsolatfelvételben, míg az archivált anyag feldolgozásában a mennyiségi problémákon tud segíteni, „cserébe” az archivált tartalom különböző módokon megjelenhet a helyi vonatkozású gyűjteményében, akár egy helyi webarchívum formájában is.¹⁹ Az ilyen jellegű együttműködés fontosságát a kezdetektől hangsúlyozta a Webarchívum, ezért távlati céljai között szerepel egy országos partnerhálózat kiépítése is, melyhez az egyik első konkrét lépést a KDS-projekt keretében történt együttműködés jelentette,²⁰ de ebbe a sorba illeszkednek a különböző képzési lehetőségek is.²¹

A Webarchívum működésének eddigi három éve alatt fokozatosan kialakultak a munkafolyamatok a különböző webhelyek archiválására, először ebben kellett megtalálni azokat a pontokat, melyekhez hasznosan, sikerrel tudnak kapcsolódni a partnerek úgy, hogy ne kelljen aránytalan terheket vállalniuk, de mégis ízelítőt kapjanak a tevékenységből és rálátásuk legyen az egészre. Szerencsére a munkafolyamatok egy része nem igényel különösebb előképzettséget, elég egy általános informatikai felhasználói készség (például a webhelyek összegyűjtése, az archivált anyagok minőség-ellenőrzése), más esetekben (például webhelyek metaadatolásánál) pedig kifejezetten előny a könyvtárosi képzettség, és mind-

ezen tevékenység hétköznapi informatikai eszközökkel végezhető. Mivel az archiválások technikai műveletei jellemzően eszközigenyesek (linuxos szerverek), ezért csak olyan lehetőség került be az opcionális feladatok közé, ami személyi számítógépeken is végezhető. Ugyanakkor célszerű volt kihasználni a helyismeretből fakadó előnyöket, és a könyvtár meggyéjére, esetleg tágabb régiójára szűkíteni a címgyűjtést,²² de érvényesülnie kellett a Webarchívum elsődleges gyűjtőkörének is.

Végül ezen szempontok mentén lett meghatározva a feladatkiírás,²³ és kezdődhetett el a munka ez év február-március folyamán.²⁴ Minden partnerünkkel egyenként felvettük a kapcsolatot videobeszélgetés keretében, részletesen ismertettük a feladatokat, az elvárásokat és átbeszéltük a munkafolyamatot is, majd ezek nyomán elkészült egy leírás,²⁵ amit elküldtünk mindenkinek. Jó döntésnek bizonyult, hogy a rendelkezésre álló idő szűkössége miatt – május végéig kellett teljesíteni a feladatokat – kettéválasztottuk a kiírásban szereplő feladatokat mindenképpen teljesítendőkre és opcionálisakra. Elsődleges feladatként azt kértük partnereinktől, hogy a megadott tematika szerint gyűjtsenek olyan helyi vonatkozású webhelyeket, amelyek még nem szerepelnek a Webarchívum gyűjteményében,²⁶ valamint válogassanak olyanokat is, melyeket érdemes lenne felvenni a nyilvánosan elérhető archívumba, és ezeken kívül a rendelkezésre álló idő és kapacitás függvényében kerüljön sor a többi feladatra.²⁷ Külön öröm volt, hogy a projektbe bedolgozó kollégák közül többeket ismerősként üdvözölhattünk, mivel már részt vettek az utóbbi években a Könyvtári Intézet akkreditált webarchiválás tanfolyamain.²⁸

Most következzen számszerűsítve is az elmúlt hónapok munkájának az eredménye.

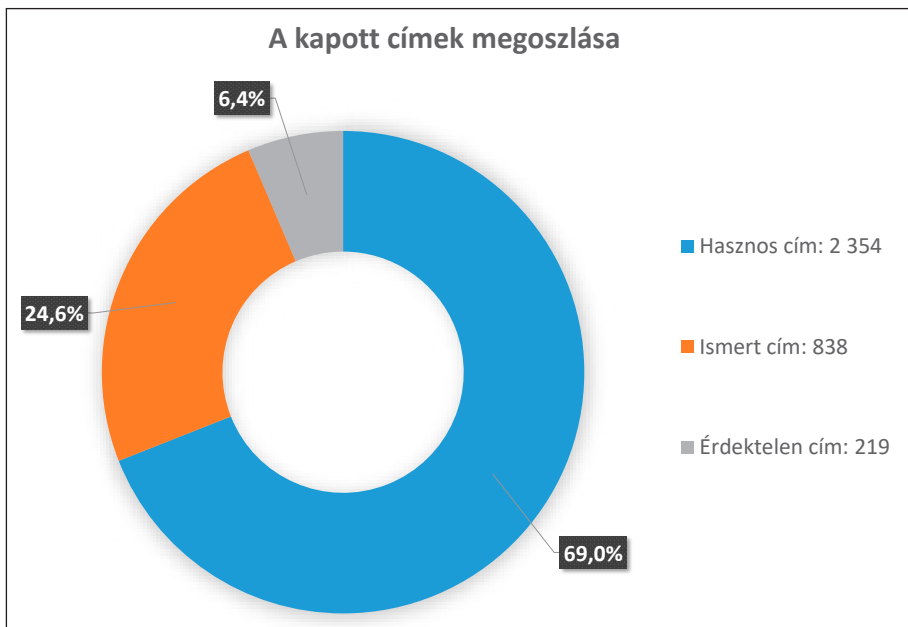
A kapott tételek darabszáma és százalékos megoszlása

Az OSZK Webarchívum jelenleg 34 ezer élő webhelyet és azok URL-jeit tartja nyilván különböző gyűjteményekben, de bizonyos tematikákhoz tartozók még nem kerültek összegyűjtésre, valamint a közösségi médiával is még éppen csak elkezdtünk foglalkozni.²⁹ Persze maga az összegyűjtés kifejezés is problémás, mert gyakorlatilag teljes körű gyűjtemény nem létezik a webhelyek változékonysága, illetve az ismeretszerzés korlátai miatt.

A pályázatban részt vevő könyvtárak nagyságrendileg ennek az állománynak körülbelül tizedére rúgó webhelyet gyűjtöttek össze³⁰ és ajánlottak a Webarchívum figyelmébe, melyből végül hasznos címként az eddigi gyűjtemény 7%-kal fog gyarapodni, ebből 2% jut a Webarchívum által korábban már feltárt tematikájú gyűjteményekbe és 5% kerül a szisztematikusan még nem feltérképezettekbe.

Június végéig – eltekintve néhány duplumtól – összesen 3411 db új tételt (plusz 31 db alternatív URL-t) kaptunk a 11 partnerünktől, melyeket egyesével ellenőriztünk ismertség, relevancia és formai szempontok (név és címalak) szerint. E munka eredményeként 3491 db besorolásra váró webhely nézett ve-

lünk farkasszemét, ugyanis tizenkettedik könyvtárként az OSZK is gyarapította saját gyűjteményét 80 db még ismeretlen tétellel (véletlenszerűen talált, például kapcsolódó címek feltárásával). A 3411 db címből az ellenőrzés során 838 db (24,6%) bizonyult általunk már ismertnek, 219 db (6,4%) érdektelennek és 2354 db (69,0%) hasznos, új címnek, amihez persze még hozzáadódik 80 db OSZK-s gyarapítás is (a bővített darabszám 2,3%-a).



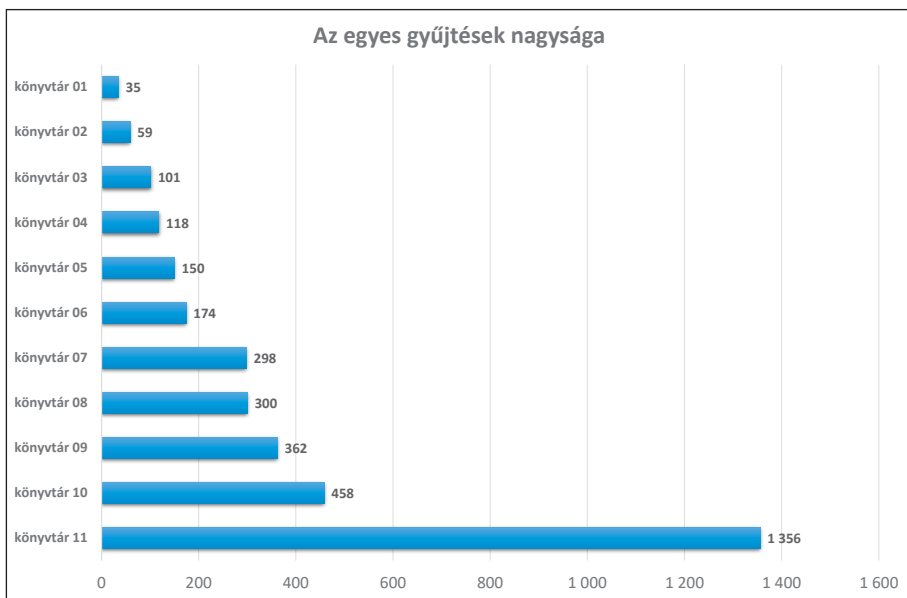
Felmerül a kérdés, miből adódtak az ismert és az érdektelen tételek? Azt minden résztvevő fel tudja mérni, mennyire tartotta magát a kért szempontokhoz (ellenőrzés, gyűjtőkör). Az ismert címek tekintetében azonban az mindenképpen árnyalja a képet, hogy a Webarchívum a projekt teljes ideje alatt nem tudta biztosítani a naprakész címlenőrzést (csak időközben készült el az új kereső³¹); illetve a véletlenek összjátéka folytán májusra jött létre az oktatási webhelyekről egy új tematikus gyűjtemény az Oktatási Hivataltól kapott lista alapján, emiatt az adatok feldolgozásakor néhány iskola honlapja már az ismert címek közé került. Az érdektelen tételek esetében az lehetett zavaró tényező, hogy az esetleg helyismereti szempontból fontos és nyilvántartandó oldal egy, a Webarchívum számára már ismert főoldal része,³² ezért nem új számára, és ez nem lett kiszűrve adatfeltöltés előtt. Szerencsére voltak olyan könyvtárak, amelyek már a gyűjtés során megmutatták addigi munkájukat, így azt észrevételezve elkerülhető vagy csökkenthető volt az esetleges hibás gyakorlat folytatása. Mindezek miatt, valamint azért, mert végül is a résztvevők számára új szempontok szerint kellett a gyűjtést végezni,

első próbálkozásra nem tekinthető rossznak az eredmény, egy esetleges újabb gyűjtés vagy folyamatos együttműködés esetén várhatóan javulna a hatékonyság.

A gyűjtések nagysága, eloszlása

Nem szeretnénk külön kitérni az együttműködésben részt vevő könyvtárak munkájának értékelésére, se darabszám, se hasznosság tekintetében. Reális értékeléshez sok egyéb szempontot is figyelembe kellene venni, azok nélkül hamis következtetések születnének egy-egy számszerűsített mutatóból – ezekről az előzetes munkamegbeszéléseken is szó volt.³³

Számszerűsítve 35, 59, 101, 118, 150, 174, 298, 300, 362, 458 és 1356 tételt kaptunk az egyes könyvtáraktól, összesen 3411 db webhely nevét, címét és esetleg egyéb adatokat. A két szélsőérték között majdnem negyvenszeres a különbség, a gyűjtések zöme a száz és ötszáz közötti darabszámból került ki. Átlagnak a 310 tétel tekinthető, ami a két szélsőérték kivonásával 224 lesz. Csak érdekességként néztük meg, hogy a legnagyobb gyűjtemény hasznosság tekintetében eltér-e az összesített átlagtól, és meglepődve tapasztaltuk, hogy szinte megegyeznek az arányok: a felvett címeknek ugyanúgy 69%-a bizonyult hasznosnak, mint az összesített gyűjtemény esetében, ismert címből viszont kevesebbet, érdektelenből többet kaptunk az összesítés átlagához képest.



A hasznos új címek tematikus megoszlása, besorolása

A projekt keretében – a partnerkönyvtárak és az OSZK révén – összegyűlt és a Webarchívum által még nem ismert hasznos címeket érdemes kettéválasztani

aszerint, hogy általunk korábban már szisztematikusan gyűjtött tematikáról van szó, vagy olyanról, mellyel kapcsolatban még nem volt érdemi címfeltárás. (Mindkét esetben az elsődleges tematika alapján került besorolásra az adott tétel és az alapján szerepel az összesítésekben.)

A Webarchívum által korábban már összeválogatott tematikus részgyűjtemények esetében 715 db új címet kaptunk, ez az eddig összegyűjtött 34 ezer webhely mintegy 2%-a, a még nem gyűjtött tematikák esetében kapott 1719 db új cím az eddigi állomány csaknem 5%-a.

Lássuk témakörök szerint a címek megoszlását az eddig ismert és a kapott új tételek szerint.

Témakör*	Nyilvántartott címek	Új cím	A nyilvántartott címek százalékában	Az új címek százalékában
Egyetemek, főiskolák	3224 db	4 db	0,12%	0,56%
Elektronikus periodikák	5232 db	53 db	1,01%	7,41%
Irodalom, irodalomtudomány és -történet	1188 db	14 db	1,17%	1,96%
Könyv- és egyéb kiadók, kereskedők	1253 db	28 db	2,23%	3,91%
Kormányzat, önkormányzatok, politikai és civil szervezetek	3609 db	93 db	2,57%	13,00%
Könyvtárak, levéltárak, múzeumok és galériák	1836 db	43 db	2,34%	6,01%
Kutatóintézetek, tudományos szervezetek	850 db	8 db	0,94%	1,12%
Média, sajtó, műsorszórás	805 db	12 db	1,49%	1,68%
Képző-, előadó-, zene- és filmművészet	7043 db	107 db	1,52%	14,96%
Kulturális intézmények, művelődési házak, rendezvényhelyszínek	660 db	38 db	5,75%	5,31%
Közoktatás és egyéb képzések	5132 db	294 db	5,73%	41,12%
Vallások, hitrendszerek, egyházak	2219 db	21 db	0,95%	2,94%
* Bizonyos esetekben nem tematikus, hanem műfaji besorolást alkalmazunk, ilyen esetekben ez az elsődleges szempont. Így kerültek külön gyűjteménybe az e-periodikák és a közösségi tartalmak (pl. Facebook-oldalak, Instagram-fiókok stb.) is. A táblázat adatai ezt tükrözik, így például egy múzeum Twitter-csatornája nem a közgyűjteményi webhelyek közé kerül.				

Érdekes látni, hogy a 12 csoportból hétben az átlag alatt maradt a gyarapodás, háromnál kicsivel, de két csoportban (kulturális intézmények, rendezvényhelyszínek) messze átlag feletti a bővülés. Utóbbiak esetében vélhetően visszaigazolódik az az előzetes feltételezés, hogy a helyismeret előnyt jelent a lokális tartalmak gyűjtésének esetében. Az oktatási webhelyek kiugró aránya tekintetében valószínűleg szerepet játszik az is, hogy ez a legfrissebb gyűjtemény, és még nem sikerült olyan szinten feltárni a vonatkozó tartalmat, mint a régebbi gyűjteményeknél (például a szakképzési centrumok esetében csak a központi intézmény szerepel a kapott nyilvántartásban, a tagiskolák nem).

A Webarchívum által eddig szisztematikusan még nem gyűjtött témakörök vonatkozásában a következő méretű gyűjtések születtek. (Ezeknél a gyűjteményeknél nem vettük figyelembe, hogy esetleg már hány címet tartunk nyilván bennük.)

Témakör*	Új cím	Összes új cím százaléka
Bölcsészet- és társadalomtudományok	2 db	0,1%
Egészségügy, szociális szféra	32 db	1,9%
Életmód, szabadidő, hobbi	22 db	1,3%
Facebook-oldalak	596 db	34,7%
Egyéb közösségi oldalak	13 db	0,8%
Idegenforgalom, vendéglátás	477 db	27,7%
Ipar, mezőgazdaság, munkaügy	141 db	8,2%
Sport, testkultúra	311 db	18,1%
Szolgáltatás, kereskedelem, közlekedés, marketing	43 db	2,5%
Természet- és műszaki tudományok	39 db	2,3%
Történelem, hely- és családtörténet	43 db	2,5%
<i>* A besorolás elve megegyezik az előző táblázatban leírtakkal.</i>		

Nem meglepő módon a legnagyobb csoport a Facebook (bár különböző mértékben lett felvéve a gyűjteményekbe), amit az idegenforgalom-vendéglátás és a sport-testkultúra követ. Érdekesen árnyalja a rangsort, hogy az idegenforgalom-vendéglátás kategória lett a legtöbbször megadva másodlagos besorolásnak, és bár ezt csak indokolt esetben alkalmaztuk, jól mutatja, mely ágazatnak van leginkább kapcsolódása más területekkel.

Az is látszik a számokból, hogy még jócskán vannak fel nem tárt webhelyek, bőven van tere a további munkának, és nem csak a Facebook esetében (az önmagában is hihetetlenül nagy és növekvő súlyú). Az ország lakosságának ugyan

kisebbség felét fedik le azok a megyék, melyek könyvtárai most részt vettek ebben a projektben, de bizonyos témák meglepően alulreprezentáltak. Ha arra gondolkunk, hogy csak állami fenntartásban 91 kórház van, de ezek mellett számtalan szakrendelő, magánkórház és -rendelő, valamint szociális intézmény van az országban, akkor a kapott 31 cím elég kevésnek tűnik. Láthatóan a sport-testkultúra tematika jobban megmozgatta a gyűjtést végző kollégákat, talán a koronavírus-járvány miatti kényszerű otthonlét okán is. Az idegenforgalom-vendéglátás ugyan nagyobb számmal szerepelt az összesítésben, de ez a tematika önmagában is hatalmas lenne, ahogy a szolgáltatás-kereskedelem-közlekedés-marketing vagy az életmód-szabadidő-hobby kategória is, amikből viszont megint nagyon kevés címet kaptunk, és sokkal több helytörténettel foglalkozó webhelyet vártunk.

Talán azt külön nem is érdemes kiemelni, hogy a helyismeret előnye visszaköszön a bővülésben, hiszen pont ez volt a cél, azt viszont már igen, hogy ennek köszönhetően kiemelt gyűjtőkori webhelyeket is megismertünk, például mindeddig „rejtőzködő” periodikákat vagy kulturális intézményeket.

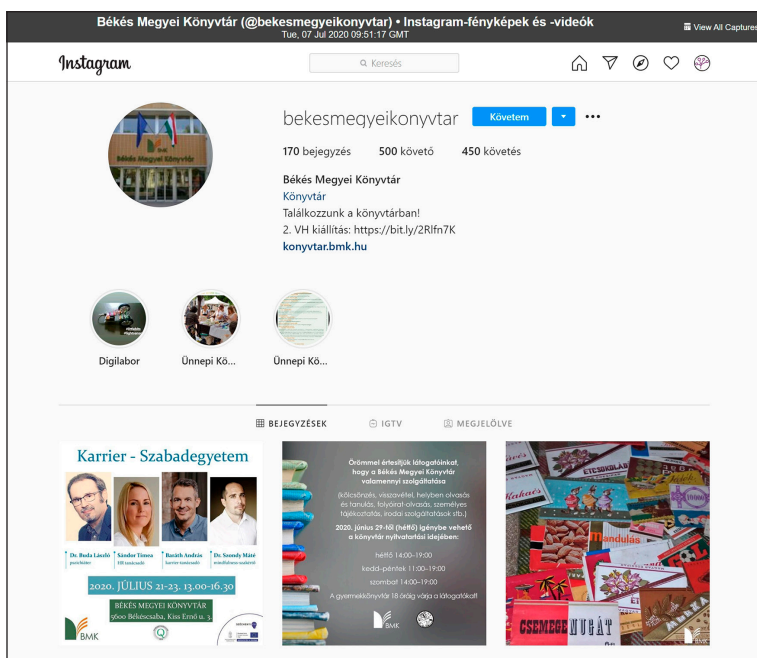
Munka után – munka előtt

A projekt során kapott webhelycímekeket időközben felvettük *tematikus részgyűjteményeinkbe* és negyedéves ütemezéssel mentéseket készítünk róluk. Ezenkívül ezek a webcímek is bekerültek a június végén indított magyar *webtáraratóba*,³⁴ mely során a tematikus gyűjteményekben szereplő címeknél nagyságrendileg bővebb körrel készítünk pillanatfelvételt.³⁵ Sajnos a közösségi média webhelyei esetében nem ilyen jó a helyzet, mert azok archiválása más jellegű technológiát igényel, és erre jelenleg nincs igazán kiforrott automatikus eljárás, ezért a legfontosabb tételek esetében manuálisan készítünk egyedi mentéseket,³⁶ a többi címet csak nyilvántartjuk.

A nyilvántartott webhelyek gyarapodása mellett, talán mégis az a legfontosabb eredmény, hogy reményeink szerint nem ért véget a közös munka. Nem ért véget, mert három résztvevővel (a kecskeméti Bács-Kiskun Megyei Katona József Könyvtárral, a tatabányai József Attila Megyei és Városi Könyvtárral és az egri Bródy Sándor Megyei és Városi Könyvtárral) már konkrét megbeszéléseket folytattunk a lehetséges további együttműködésről, és más könyvtárak is jelezték, hogy szívesen folytatnák a közös munkát. Az együttműködés szorosabbra fűzése érdekében elkezdtünk előkészíteni egy megállapodástervezetet is a jövőbeli közös munka lehetséges kereteiről, amelyről minden érintettel egyeztetni fogunk. Örömmel láttuk, hogy sokan felismerték az élő és az archivált webtartalom fontosságát, helyét a saját, helyi gyűjteményeinkben.

Nem ért véget a közös munka, mert a java még hátra van. Egyrészt a webtartalom változékonysága megköveteli a most létrehozott állomány folyamatos gondozását, hogy a közös eredmény ne enyvésszen el idővel, ugyanis hiába gyarapodott most a Webarchívum sok új és értékes webhellyel, hiába kerülnek be ezek

a tételek a helyi gyűjteménybe, ha a ma adatai infálódnak az idő múlásával. Másrészt célszerű kiterjeszteni az együttműködők körét újabb partnerek bevonásával, hogy finomítani és bővíteni lehessen az elért eredményeket a közös feladat, a jelen webes kultúrájának megőrzésének érdekében. Távlatilag szeretnénk a többi megyei hatókörű városi könyvtárral, valamint a szak- és egyetemi könyvtárakkal is felvenni a kapcsolatot és a mostanihoz hasonló együttműködést kialakítani, ahogy más, a webarchiválásra nyitott közgyűjteménnyel is, és persze ne feledkezünk el a határon túli magyar kulturális intézményekről sem!



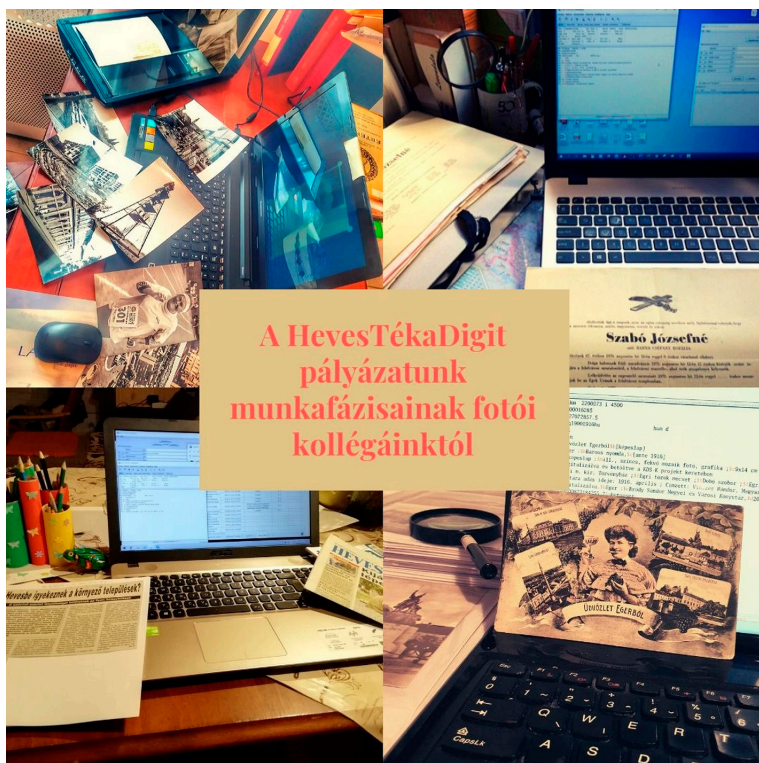
Egy frissen lementett Instagram-oldal archivált változatának részlete

A KDS-projekt keretében végzett közös munkáról az immár rendszeres őszi workshopunkon³⁷ is szeretnénk számot adni. A rendezvény tervezett témája az intézményi együttműködés lesz, és az előadások mellett – melyre a mostani partnereink közül is szeretnénk felkérni előadót, aki a saját és intézménye szemszögéből értékelné a közös munkát – nagyobb hangsúlyt kapnának a gyakorlati szempontok.

Sokféle számadat szerepelt ebben az összefoglalóban, azonban ha meg szeretnénk vonni az együttműködés mérlegét, mégsem a Webarchívumban lévő gyűjtemény számszerű gyarapodását emeljük ki, hanem azt, hogy több nagy, megyei hatókörű városi könyvtár és sok könyvtáros kolléga kapott ízelítőt a webarchiválás feladataiból, nehézségeiből, és láthatóan sokukban alakult ki ér-

ben a Bródy Sándor Könyvtár munkatársai is bekapcsolódtak a helyi, regionális vonatkozású webhelyek címeinek a gyűjtésébe.

A 2020. április 27-ei összdolgozói értekezleten – a járványhelyzet miatt a Zoom alkalmazás segítségével – rövid előadás hangzott el az OSZK webarchiválási tevékenységéről, az eddig elért eredményekről, a pályázati kiírásban megfogalmazott feladatokról és annak megvalósításának lehetőségeiről. A tájékoztató után felhívást tettünk közzé a könyvtár közösségi oldalain³⁸ és sajtóközleményben tájékoztattuk a helyi média³⁹ munkatársait a gyűjtésről. Külön levelet küldtünk a Heves megyei városi könyvtárak és a könyvtári szolgáltató helyek munkatársainak is, hiszen ők ismerik leginkább a megyei településeken élő művészeket, alkotókat.



Otthon is dolgoztunk – home office

A felhívásunkra értékes helyi vonatkozású oktatási, tudományos és kulturális témájú online tartalmak érkeztek, volt olyan könyvtári szolgáltató helyen dolgozó könyvtáros kolléga, aki már tematikusan küldte a webhelyek címeit. Az egy hónapig tartó gyűjtés alatt természetesen sok olyan webhelyet kaptunk vagy találtunk, amelyek már szerepeltek az OSZK gyűjteményében – így történt az oktatási intézmények elérhetőségeivel is –, szerencsére a keresést megkönnyítette a webarchívum részgyűjteményi keresője. Több esetben előfordult, hogy egy

14. A KDS-K aloldala az oszk.hu-n <http://www.oszk.hu/kds-k>; http://www.oszk.hu/kds-k/palyazat_2019 (2020.08.25.); Pályázat a Közgyűteményi Digitalizálási Stratégia keretében a könyvtári digitalizálás támogatására (továbbiakban Pályázati felhívás): http://www.oszk.hu/sites/default/files/KDS-K-2019_P%C3%A1ly%C3%A1zati-felhiv%C3%A1s_modosított.pdf (2020.08.25.)
15. Pályázati felhívás 5. p.
16. Lásd: Drótos László – Visky Ákos László: *Rákóczi-archívum*. In: Könyv, Könyvtár, Könyvtáros, 29. évf. 2020. 3. sz. 36. p.; https://epa.oszk.hu/01300/01367/00326/pdf/EPA01367_3K_2020_03_035-048.pdf (2020.08.25.)
17. Ezt egy eseményalapú archívum kialakításával terveztük megvalósítani, melynek apropóját a II. Rákóczi Ferenc emlékévk adta. Az azóta létrejött honlap: <https://rakoczi2019.webarchivum.oszk.hu/> (2020.08.25.)
18. Pályázati eredmény a könyvtári digitalizálás támogatására: <https://www.kormany.hu/hu/emberi-eroforrasok-miniszteriuma/kulturaert-felelos-allamtitkarsag/hirek/palyazati-eredmeny-a-koonyvtari-digitalizalas-tamogatasara> (2020.08.25.)
19. Kokas Károly a 2019-es 404 Not Found – Ki őrzi meg az internetet? workshopon elhangzott előadásában ennek egy lehetséges módját mutatta be. A prezentáció megtekintése mellett érdemes meghallgatni az előadás szövegét is: <https://prezi.com/4-8moxcxoe7/szegedikum-a-webarchivumban/> (2020.08.25.); <http://videotorium.hu/hu/recordings/35069/04-kokas-karoly-szegedikum-a-webarchivumban> (2020.08.25.);
20. <https://webarchivum.oszk.hu/szakembereknek/egyuttmukodesi-lehetosegek/> (2020.08.25.)
21. <https://webarchivum.oszk.hu/szakembereknek/tanfolyam-es-e-learning/> (2020.08.25.)
22. Bármilyen hihetetlennek tűnik, hogy akkor, amikor interneten keresztül a világ bármely pontjára eljut az ember, egy szisztematikus címfeltárás során milyen korlátokba lehet ütközni (például, amit nem „lát”, nem tesz élénk a Google-kereső, az szinte nincs is), és mekkora előnyt jelentenek a lokális kapcsolódások, illetve az a hagyományos munka, amit a helytörténeti vagy más szakosított gyűjteményekbe fektetnek a könyvtárak és egyéb intézmények. Az, hogy mennyire működött ez a KDS-projekt esetében, reméljük majd kiderül az összefoglalóból.
23. Lásd korábban a hivatkozott Pályázati felhívást.
24. Színes adalék a projekt történetéhez, hogy még normál, hétköznapi keretek között kezdtük el a munkát, de március közepétől a koronavírus okozta járványhelyzet miatt mindenki rendkívüli körülmények közé került és valószínűleg nem szokványos kihívásokkal szembesült. Bár arról nem rendelkezünk pontos információkkal, hogy partnereinket miként érintette ez, mennyire segítették, netán hátráltatták a körülmények a feladatok teljesítését – és ne feledkezzünk meg arról sem, hogy a projekt keretében más jellegű vállalásokat is teljesíteniük kellett –, hozzánk nem érkezett kifogás emiatt, az összegyűjtött webhelyek átadása megtörtént.
25. Visky Ákos László: *Feladatok a KDS-K-projektben a webarchiválás támogatására – kerítariumok, követelmények*. https://webarchivum.oszk.hu/wp-content/uploads/2020/07/Feladatok_a_KDS-K-projektben_a_webarchivalas_tamogatasara.pdf (2020.08.25.)
26. Egy megosztott Google-táblázatba kellett feltölteni a tételeket, adatként a webhely nevét, URL-jét, esetlegesen kapcsolati információt és egyéb fontosnak vélt megjegyzést kértünk rögzíteni. Az ismertséget kezdetben egy – sajnos nem naprakész – űrlappal, majd később az időközben elkészült új webarchívumhonlap keresőjével lehetett ellenőrizni.
27. Több részfeladatot is kínáltunk előzetesen, meghagyva a választás lehetőségét. Optimális esetben az összegyűjtött webhelyek archivált változatait és oldalképet lehetett volna ellenőrizni minőség szempontjából, valamint ki lehetett volna próbálni metaadatok rögzítését is. Később még felvetettük az egyedi mentések készítésének lehetőségét is a Webrecorder vagy HTTrack nevű szoftverekkel.

28. <https://ki.oszk.hu/tanfolyamok/az-internet-archivalasa-mint-kozgyujtemenyi-feladat> (2020.08.25.)
 29. <https://webarchivum.oszk.hu/webarchivum/reszgyujtemenyek-szerint/tematikus-aratasok/> (2020.08.25.)
 30. A Webarchívum eddigi működésének három éve alatt az egyéb feladatok mellett nem folyamatosan gyűjtötte a címeket és gyakorlatilag érintetlen terepen mozoghatott. A projekt ideje alatt a partnerkönyvtáraknak három hónap állt rendelkezésükre úgy, hogy arra is figyelniük kellett, hogy ismert címet ne vegyenek fel a gyűjtésükbe.
 31. <https://webarchivum.oszk.hu/webarchivum/kereses/a-reszgyujtemenyekben/> (2020.08.25.)
 32. Például megkaptuk a megyei levéltár honlapját, ami ma már a központi honlap egyik aloldala (alkönyvtár a struktúrában) és nem önálló webhely, mint korábban volt. Mivel a webhelyeket az aratások technikai feltételei (mi minősül önálló webcímnnek, vagyis főszabály szerint csak a domaineket és aldomaineket) szerint tartjuk nyilván, ezek az adatok érdektelenek számunkra, mert az arató robot a kezdőoldaltól kiindulva végigmegy a struktúrában és eljut az aloldalig is. Alkönyvtári címet kivételként csak akkor tartunk nyilván, ha szükséges az adott tartomány eléréséhez, esetleg ott külön kezelt, kiemelt tartalom van, például e-periodika stb.
 33. Mivel nem ismert az egy-egy térséghez köthető internetes tartalmak gazdagsága, és az sem, hogy ezek közül mennyi ismert már a Webarchívum számára, ezért például mennyiségi paraméter sem volt megszabható, hogy hány darab új webhely összegyűjtését várjuk el minimumként.
 34. <https://webarchivum.oszk.hu/blog/2020/07/06/idei-also-webter-aratas/> (2020.08.25.)
 35. A webtartalom archiválása jellemzően különböző szempontok szerint létrehozott gyűjteményeken keresztül, de eltérő paraméterekkel történik. Például negyedévente aratjuk a válogatott és gondozott tematikus és az elektronikus periodikagyűjteményeket, melyek esetében részletesebb archiválásra törekszünk, ezért az egyes webhelyekről nagyobb mélységből is próbáljuk az anyagokat lementeni (jellemzően 3-5 szint). Míg a magyar webteret félévente megcélzó, hozzávetőleg tízszer több webhelyet bejáró archiválás esetén az a szempont, hogy minél nagyobb körből, időben koherens, de egy-egy webhelyről kevesebb anyag kerüljön begyűjtésre az általános webtartalom reprezentálása céljából (ez esetben csak két szint mélyre mehet a robot). Mint látható, egyik esetben sem várható egy webhely teljes körű, hanem csak részleges archiválása és megőrzése, aminek különböző okai vannak – de annak kifejtése már egy másik cikk témája lenne.
 36. 2020 tavaszán kezdtünk el kísérletezni a közösségimédia-tartalmak archiválásával, ezidáig leginkább az Instagram-oldalak mentésével foglalkoztunk. Ezen webhelyek jellegzetessége, hogy a felhasználó aktivitását követve töltődnek be a tartalmak, amit egy „vak” aratórobot nem tud lekezelni. Ezeknél az oldalaknál a Webrecorder nevű szoftvert használjuk archiválásra, de ez nem automatikus művelet, hanem valakinek végig kell kattintgatni az adott oldalt, tulajdonképpen böngészni rajta, amit rögzít a program és szabványos archív formátumban elment. Sajnos nem csak emiatt nehézkes a használata, ugyanis nehezen birkózik meg ezen webhelyek folyamatosan változó megoldásaival.
 37. <https://webarchivum.oszk.hu/szakembereknek/404-not-found-workshop/> (2020.08.25.)
 38. <https://www.facebook.com/brodykonyvtar/photos/a.124134777656553/3571456436257686/?type=3&theater> (2020.08.25.)
 39. <https://www.egriprogramok.hu/hu/kulturalis-hirek/4868-hevestekadigit-palyazat-azaz-digitalizal-a-brody-sandorkoenyvtar.html> (2020.08.25.)
-