

Doktori (PhD) értekezés tézisei

**A webarchiválás elméletének és gyakorlatának
alapelemei**

**A szervezett keretek között zajló webarchiválás
kezdetei Magyarországon**

Németh Márton

Témavezető: Eszenyiné dr. Borbély Mária



DEBRECENI EGYETEM

Informatikai Tudományok Doktori Iskola

Debrecen, 2021

A kutatás tárgya, az annak alapjául szolgáló hipotézisek

A Phd dolgozat alapjául szolgáló kutatás a webarchiválás elméleti alapelemeinek bemutatása mellett annak első lépéseit is fel kívánja villantani, hogy az elméleti kereteket miként lehet magyar körülmények között a gyakorlat nyelvére lefordítani, az Országos Széchényi Könyvtárban 2017-ben induló webarchiválási projekt keretei között. A kezdeti kutatási hipotéziseket az alábbiakban öt fő szempont szerint lehet megfogalmazni:

1. A kutatás fő hipotézise, hogy magyar nyelven, az újdonság erejét is felvillantva áttekintő elemzés nyújtható a webarchiválás létjogosultságáról, elméleti háttéréről, értelmezési tartományáról, elméleti és gyakorlati korlátairól, közgyűjteményi kereteiről, kutatási célú felhasználásának módozatairól, s emellett feltárhatók a témakör nemzetközi gyakorlatának, illetve szakirodalmi reprezentációjának legfontosabb vonásai is.

2. Fontos kérdésként merült fel kezdeti hipotézisként, hogy az elméleti és a nemzetközi gyakorlati környezetek felvillantása mellett, megteremthető az Országos Széchényi Könyvtár keretei között a szervezett webarchiválás archiválási és szolgáltatási gyakorlata, a megfelelő munkafolyamatok kialakításával, a megfelelő pénzügyi, információtechnológiai, illetve jogi háttér megteremtésével. Amennyiben nem, vagy csak részlegesen kerül erre sor, melyek a fő kihívást jelentő tényezők?

3. Nélkülözhetetlennek tűnik annak tisztázása is, hogy milyen lényeges vonásai vannak a mostanában kialakuló webkönyvtárosi, webkurátori munkaköröknek, s meddig terjed ezen tevékenységek határa a könyvtári keretek között, s mi tartozik a többi közgyűjteményi ág, illetve az informatika keretei közé.

4. Az esetleg felbukkanó szakmai és pénzügyi, információtechnológiai infrastrukturális nehézségek, a jogi, szabályozási háttér megteremtésének nehézkes folyamata nem lehetetlenítik-e el, hogy egy kutatás-fejlesztési kísérleti projektből állandó nemzeti könyvtári szolgáltatás válhasson?

5. Meg lehet-e teremteni a webarchiválás témakörének bevonását az akkreditált felsőfokú könyvtáros továbbképzés, illetve a felsőoktatási, egyetemi képzés keretei közé?

A hipotézisekre adott válaszok egyes szakmai elemei a dolgozatba szervesen beágyazva, illetve a kutatás eredményeit összegző tézisek formájában kerülnek kifejtésre, az összkép felvázolására pedig a dolgozat zárófejezetében kerül sor. A dolgozat további egyedi vonását adja meg, hogy az egyéni kutatási program megvalósítása a hipotézisekben megfogalmazottak megválaszolása során néhány ponton keveredik azokkal a kollektív erőfeszítésekkel, melyeket az OSZK E-Könyvtári Szolgáltatási, majd 2020 tavaszától önálló Webarchiválási Osztálya keretében munkatársaimmal együtt tettünk az üzemszerű webarchiválási szolgáltatás feltételeinek megteremtéséért. A dolgozat szövegében az egyes szám első személy használata mindig olyan tartalmi elemre utal, mely az egyéni kutatáshoz, elemzéshez, kutatási eredmények feltárásához kapcsolódik. A témához tartozó kollektív, gyakorlati erőfeszítések lenyomatait pedig a többes szám első személy használata tükrözi vissza. Az egyéni kutatás gyakorlati elemeinek, illetve a kollektív gyakorlati tevékenységek eredményeinek felvillantására több fejezetben is kitérek, az összegzést pedig ezen a téren is a záró fejezetben végzem el.

Tézisek

[T1.] Kulturális örökségünknek egyre nagyobb szelete a számítógépes világhálón tűnik fel. Ennek begyűjtése, feldolgozása, hosszútávú megőrzése összetett kihívást jelent. E kihívás kezelése számos fontos összetevővel bír az információtudományi, informatikai, társadalomtudományi és bölcsészeti tudományos területeken.

1.1 A webarchiválás létjogosultsága

A webarchiválás szükségszerűsége információtudományi szempontból sokáig megkérdőjeleződött a számítógépes világhálóról való gondolkodás fényében. Felvetődött, hogy az egész világhálót nem lehet-e valamiféle hatalmas digitális könyvtárként, tudásszervező rendszerként felfogni, amit nem is kell archiválni,

mert a rendszeren belül intakt módon megőrződnek a tartalmak. Mint ahogy arra Niels Ole Finnemann információkutató rámutat, ez a megközelítés több okból is problémás (Finnemann, 2019). Először is, a web egyszerűen túl nagy ahhoz, hogy bármilyen tartalomfeltáró, tartalomkezelő eszközzel átfogható, kezelhető legyen. A második fő kihívás a web állandó jelenidejűsége, efemer jellege. A tartalmak nagyon gyors ütemben, követhetetlen sebességgel jelennek meg és tűnnek el a világhálóról. Már a szervezett webarchiválás kezdeteinek idején, 1996-ban is csak 42 napra becsülték egy honlap átlagos élettartalmát, s ez azóta nemhogy növekedett volna, hanem radikálisan le is csökkent. Emellett a webes tartalmak jó részének nincs megfogható szerkesztője, s a sok esetben az elvégzett tartalmi változtatások is követhetetlenek. A digitális világ ebben az értelemben radikálisan véget vetett az információhordozó objektum stabilitásának. A harmadik fő indok szerzői oldalról vet fel kihívásokat. A világhálón nincs semmiféle kapuőri szerep, bárki bármit publikálhat. Nagyon magas fokú heterogenitás nem képezhető le semmiféle hagyományos közgyűjteményekben használatos taxonómiai eszköztárral. Az áttekinthetetlen összetettséghez ez a jellemző is nagymértékben hozzájárul (Finnemann, 2019).

A webarchiválás tudományos és társadalmi jelentősége ellenére mind a tudományos és társadalmi kommunikációban, mind pedig a kulturális örökséghez kötődő pénzügyi erőforrások elosztásakor globális szinten is csak csekély figyelmet kap. A nemzetközi tudományos párbeszéd e szakmai területen csupán az utóbbi öt-tíz évben kapott igazán lendületet.

1.2 A Webarchiválás értelmezési tartománya

A webarchiválás tudományos értelmezési kereteinek kijelölésekor első alapvető állításom, hogy a számítógépes világháló mint rendkívül összetett és strukturált hálózat a maga egyediségében nem hasonlítható össze igazán egyetlen más olyan hagyományos vagy digitális információforrással sem, amellyel az emberiség történetében eddig találkozhattunk. A második fő állításom, hogy a legnagyobb kihívás a közgyűjtemények, illetve bármely más archiválást végző intézmény, magánszemély számára, hogy a világháló annak eredeti természetes szerkezetében, összetettségében, folyamatosan változó dinamikus tartalmi kínálatával, a maga teljességében nem archiválható. A web élő szövetéből

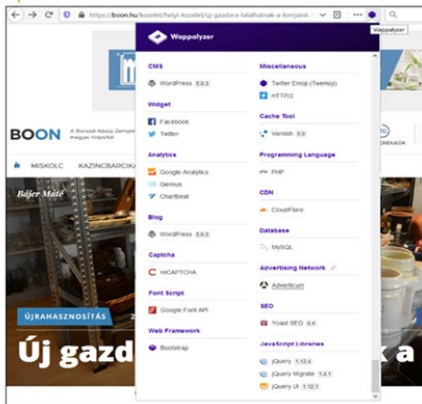
archiválási céllal csupán önkényes töredékeket (egyres webhelyeket, honlapokat, részben vagy egészben, illetve azok meghatározott gyűjteményét) tudunk kiragadni és megőrizni, ami persze a maga töredékességében is jelentős fontossággal bír. A válogatott tartalom elemeit körülvevő teljes eredeti kontextus teljes megőrzésének igénye azonban illúzió csupán. A webes archiválás ezen alapvető sajátosságát mindig észben kell tartanunk, ha bármilyen vonatkozásban szóba kerül ez a tudományos terület. A harmadik állítás az utóbbiakból következők: a webarchiválás a maga egyre erősebben automatizált formájában is alapvetően szubjektív emberi tevékenység. Az egységes egészből a töredékek kiszakításának formáját, szempontjait, tartalmi sajátosságait ugyanis az archiválási stratégiát alkotó szakemberek határozzák meg. Ez alapján alkotják meg, illetve módosítják a begyűjtésben segítő szoftverek algoritmusait is.

A webarchiválásnak tudományos témaként történő megjelenését eleve a technikai fejlődés teremtette meg, s ez a folyamatosan zajló műszaki-társadalmi evolúciós folyamat alapvetően befolyásolja a kutatás tárgyának sorsát is.

A webszervereken napjainkban már mindenféle scriptek és egyéb szoftverek futnak. A böngészőprogramok bonyolultsága szinte már elérte az operációs rendszerek szintjét. Elterjedtek az olyan megoldások is, mint a dynamic web page („Dynamic web page”, 2020), a dynamic HTML (DHTML) („Dynamic HTML”, 2020) és a Rich Internet Application vagy Rich Web Application (RIA) („Rich web application”, 2020). Közös jellemzőjük, hogy a korai statikus, egyszerű szerkezetű weboldalak helyett szerver és/vagy kliens oldalon futó parancsokat és programokat futtatnak, használnak. Valójában már önálló alkalmazásnak számítanak (pl. közösségi oldalak, webáruházak, internetes bankok, online játékok fejlesztésének terén).

Ráadásul, amíg a számítógépes világháló eredetileg alapvetően pc-s böngészőn keresztüli tartalomszolgáltatási és hozzáférési módként jelent meg, mára már külön speciális (mobil) platformok (operációs rendszerek) jelennek meg („Mobile operating system”, 2020). A különféle platformokon olyan plusz kényelmi funkciók érhetők el, például közösségi média alkalmazásokban, melyek lényegesen más felhasználói élményt nyújtanak, mint a komputeres változatok. Az emberek többsége egyre inkább ezeken a mobilplatformokon keresztül fér hozzá a webhez. A különféle platformokra specializált felbontású tartalmak

A web egy hipermédia dokumentum?



CMS: WordPress 5.0.3

Widget: Facebook, Twitter

Analytics: Google Analytics, Gemius, Chartbeat

Blog: WordPress 5.0.3

Captcha: reCAPTCHA

Font Script: Google Font API

Web Framework: Bootstrap

Miscellaneous: Twitter Emoji (Twemoji), HTTP2

Cache Tool: Varnish5.0

Programming Language: PHP

CDN: CloudFlare

Database: MySQL

Advertising Network: Adverticum

SEO: Yoast SEO 5.8

JavaScript Libraries: jQuery 1.12.4, jQuery Migrate 1.4.1, jQuery UI 1.12.1

2. ábra A hírportál szolgáltatás mögött álló informatikai technológia (forrás: Drótos László: *Az OSZK webarchívumának újdonságai*.

http://mekosztaly.oszk.hu/mia/doc/workshop/Drotos_Laszlo_Az_OSZK_webarchivumanak_ujdonsagai.pptx)

Tovább súlyosbítja a helyzetet, hogy a weben használt formátumok, szoftvertechnológiai elemek jó része gyorsan elavul. A webarchiválás során a tartalom mellett azt a szoftveres környezetet is archiválni kell, amelynek környezetében az megjelenik, hogy később emulálható módon megjeleníthetővé váljon.

Weboldalakból sok-sok milliárd van a felszíni weben, vagyis a Google által is használt programrobotokkal bejárható nyilvános szervereken, és ezek többsége ráadásul változik (akár naponta többször is). A következő fő állításom tehát, amire a kutatás ezen aspektusa során jutottam, hogy a web jelenleg a legnehezebben megőrizhető médium. Másfelől viszont arra is utaltam már, hogy mindemmellett egyben napjaink legfontosabb információhordozó eszköze is egyben (Drótos, 2018b). Tehát az összefoglaló állítást úgy tudom megfogalmazni, hogy ha töredékesen és pillanatfelvételeket készítve is, de mentéseket kell a világhálóról készíteni akár jelenlegi, akár jövőbeli felhasználási lehetőségek, kutatások számára.

Kapcsolódó előadások, publikációk

- Drótos, L., & Németh, M. Az OSZK-ban folyó kísérleti webarchiválási projekt első évének tapasztalatai. Tudományos és Műszaki Tájékoztatás, 65(7–8), 389–400, 2018
- Németh, M.: Webarchívum mint a tudományos kutatások tárgya. Tudományos és Műszaki Tájékoztatás 67 (12), 757-765, 2020
- Németh, M., Drótos, L.: Web Archives as a Research Subject. In: Information and technology transforming lives: connection, interaction, innovation: Proceedings of the XXVII. BOBCATSS Symposium. Ed.: Gordana Gašo, Mirna Gilman Ranogajec, Jure Žilić, Madeleine Lundman, University of Osijek, Osijek, 471-478, 2019. ISBN: 9789533141213
- Németh, M., Drótos, L.: Web-Archiving: responsibility of Public Collections in the Preservation of the Digital Culture. In: Bobcatss 2018 Riga The power of reading: proceedings. Eds.: Lelde Petrovska, Baiba Ivane-Kronberga, Zane Meldere, University of Latvia Press, Riga, 124-126, 2018. ISBN: 9789934183539

[T2] Amennyiben a mindennapi élet különféle kereteinek webes lenyomatait a közgyűjteményekben nem tudjuk megőrizni és bemutatni, ezen intézményrendszer fokozatosan képtelenné válik a kultúra átörökítésére, a kortárs társadalmi problémákra történő reflexió képességét is elveszíti. Végző soron tehát a legitimitása kérdőjeleződik meg. A webarchiválással tehát igen súlyos társadalmi jelentősége miatt is tudományos szinten kell foglalkozni.

2.1 A webarchiválás társadalmi környezetének alapvonásai

Joggal vetődhet fel kérdésként, hogy miért kell a webarchiválásnak kutatási témaként megjelennie. A webarchiválás mint tudományos téma előtérbe kerülése számos tényezővel indokolható. Megváltoznak a kulturális örökségünk áthagyományozásának keretei. A közgyűjtemények által őrzött hagyományos írásbeli és tárgyi kultúrkincs jól meghatározható dokumentumhordozókhoz kötött, megfelelő infrastrukturális körülmények biztosításával ezek hosszútávú megőrzése jó eséllyel garantálható. Kulturális örökségünknek egyre nagyobb szelete azonban manapság már a számítógépes világhálón tűnik fel. Ennek a szegmensnek a begyűjtése, feldolgozása, hosszútávú megőrzése kézenfekvőnek tűnik, másfelől viszont igen összetett kihívást jelent.

A web története mintegy harminc évre nyúlik vissza. A svájci CERN intézetben eredetileg a hipertext és az internet összeházasításának fő célja a kutatási anyagokhoz való hozzáférés segítése volt. Harminc év alatt nagyon hosszú utat tettünk meg. A számítógépes világháló az emberi tudás tárolásának és áthagyományozásának alapvető médiumává lett. Amennyiben a kortárs kultúra lenyomatait a közgyűjteményekben nem tudjuk megőrizni és bemutatni, ezen intézményrendszer gyorsuló ütemben veszítheti el a kultúra átörökítésének képességét, a kortárs társadalmi problémákra történő reflexió képességét. Végző soron tehát a legitimitása kérdőjeleződik meg. A webarchiválással tehát igen súlyos társadalmi jelentősége miatt, tudományos szinten is foglalkozni kell. Fontos feladatunk elkerülni, hogy az utókor, mint valamiféle digitális középkort tartsa majd számon ezt a történeti kort, melyben élünk, miután az ebben keletkezett információk jó része nem öröklődik át a számukra.

A webarchiválás társadalmi jelentőségének felismerése még a nemzetközi szinten is várat magára. Számos részlettel szolgállok a webarchiválás közgyűjteményi kereteit ismertető első fejezet bevezetésében a terület alulfinanszírozottságáról, arról a paradox helyzetről, miszerint a tartalomelőállításra folyósított költségek elenyésző filléres hányadát sem költik el a tartalom megőrzésére. Axiómaként jelenik meg Rosenthal (Rosenthal 2017) tanulmányában, hogy megfelelő pénzügyi erőforrások biztosításával nem lenne probléma az archiválással. Világszinten 30 millió dollárt költenek azon tartalmak archiválására, melyek előállítási költsége százmilliárd dollárokban mérhető. A megőrzési szempont a tartalomelőállítás során többnyire egyáltalán fel sem merül, vagy teljesen háttérbe szorul. Lényeges állítás számomra a szakirodalom tanulmányozása alapján, (Brunelle, Kelly, Weigle, & Nelson, 2016) hogy a hatékonyabb szolgáltatás oldalán az archiválás igényei jelenleg tartalomszolgáltatói szinten bármikor, bárki számára feláldozhatók. A tartalomszolgáltatások költségét alapvetően előfizetésekben, illetve hirdetésekben fedezik. Mindkét bevételi oldal gyenge lábakon áll jelenleg a hirdetésblokkoló programok elterjedt használata, illetve az előfizetési hajlandóság csekély volta miatt. Erre válaszul a tartalomszolgáltató lobbis hatalmas nyomás alá helyezte a W3C szabványalkotó testületet, hogy a digitális másolásvédelem integrálódjon bele a html szabványba, illetve másik oldalról szabványszinten akadályozzák a

hirdetésblokkoló programok használatát. Ezek a törekvések, ha megvalósulnának, akkor a számítógépes világháló jó részét egyszerűen használhatatlanná tennék (Rosenthal, 2017). Miben segíthetünk egyszerű állampolgári szinten a webarchiválás ügyén? Erre Rosenthal szerint több mód is adódik. Az Internet Archive Save Page Now szolgáltatásával saját magunk kezdeményezhetjük az általunk hasznosnak tartott oldalak mentését. Az Internet Archive működését pénzügyi adománnyal, illetve tartalmak archiválásra történő felajánlásával is lehet segíteni. Győződjünk meg arról, hogy az adott ország nemzeti könyvtára fontosnak tartja a világháló archiválását, aktív tevékenységet folytat e téren, s támogassuk erőfeszítéseiket (nálunk az OSZK kísérleti projektje révén vált önálló szolgáltatássá napjainkban ez a tevékenység). Civil összefogással széles körű tiltakozás megszervezésével akadályozzuk meg, hogy bármiféle W3C szabványba bekerüljön a digitális másolásvédelem kiterjesztése (Rosenthal, 2017). A tanulmány megállapításai kiegészítődnek még egyéb szerzői jogi kihívásokkal is, mint például annak a veszélyével, hogy az európai szellemi tulajdonjogi szabályozás reformja elérhetetlenné teszi az amerikai Internet Archiveban a webről archivált állomány egy lényeges hányadát, vagy akár a teljességét, mivel a szerzői jogi státusza a közzétett anyagoknak számos esetben nem tisztázott (Rosenthal, 2017). A kutatásnak ebből az eleméből tehát az a fontos állítás adódik, hogy a jogi szabályozás is alapvető jelentőségű a webarchiválás közgyűjteményi megalapozásának szempontjából. Ezt a témakört önálló alfejezetben tárgyalom a dolgozat keretei között.

Végezetül megállapíthatjuk, hogy a társadalmi kontextus jelenleg nem tűnik túl biztatónak, de a jelen disszertáció elkészítésének az egyik legfőbb indoka éppen az volt, hogy ráirányítsa a figyelmet a webarchiválás területére s annak fontosságára.

Kapcsolódó előadások, publikációk

- Németh, M., Drótos, L.: Web-Archiving: responsibility of Public Collections in the Preservation of the Digital Culture. In: Bobcatss 2018 Riga: The power of reading: proceedings. Eds.: Lelde Petrovska, Baiba Ivane-Kronberga, Zane Meldere, University of Latvia Press, Riga, 124-126, 2018. ISBN: 9789934183539

- Drótos, L. & Németh, M. Beszámoló az OSZK webarchiválási pilot projektjéről, valamint a nemzetközi tapasztalatokról. OSZK Könyvtártudományi Szakkönyvtár, Budapest, 2017.
- Németh, M. 404 Not Found – Ki őrzi meg az internetet: Webarchiválás workshop az Országos Széchényi Könyvtárban. Tudományos és Műszaki Tájékoztatás, 64(11), 577–582, 2017.
- Németh, M. (2017). Nemzetközi körkép a webarchiválás gyakorlatáról. Könyvtári Figyelő, 63(4), 575–582.

[T3] Magyarországon jelentős időbeli késéssel indult el a szervezett keretek között zajló webarchiválás a nyugat -európai centrumországokhoz képest. A webarchiváláshoz kapcsolódó tudományos kihívások felvázolása, a nemzetközi kutatás-fejlesztési eredmények implementálása a magyar webarchiválás gyakorlati tevékenységi keretei közé időszerű feladatot jelent. A korszerű archiválási környezet megalapozásával jelentősen csökkenthető a lemaradás következményeinek negatív hatása.

3.1 Az internet archiválás nemzetközi környezete

A webarchiváláshoz kötődő tevékenységek és elméleti kutatási irányok, keretek közgyűjteményi szempontú feldolgozására, a legújabb nemzetközi kutatási eredmények bemutatására, ötvözve a saját itthoni kutatói tapasztalatokkal, önálló tudományos munka formájában magyar nyelven még nem került sor. Bármely, a területet érintő résztémakör mélyebb jellegű (magyar tudományos keretek között zajló) tudományos megtárgyalásához előfeltételként szükségesnek érzem egy általános bemutató jellegű elemzés elvégzését. A téma jelentősége a kulturális örökség áthagyományozásának szempontjából véleményem szerint feltétlenül indokolja tehát egy magyar nyelvű, saját kutatási eredményekre is támaszkodó, áttekintő jellegű tudományos munka elkészítését e témakörben. A Magyar Internet Archivum pilot projektjének formálódása közben, egyben a saját webarchiválással kapcsolatos kutatások megalapozásaként is elsődleges feladat volt feltérképezni az internet archiválás főbb nemzetközi szereplőit, és tanulmányozni a különböző nemzeti webarchiválási modelleket a hasznosítható jó gyakorlatok szempontjából. A projekt előrehaladása közben az OSZK csatlakozni tudott a webarchiválásban érdekelt szereplőket összefogó

International Internet Preservation Consortiumhoz (IIPC, 2019), ennek köszönhetően aktívan bekapcsolódtam több munkacsoport tevékenységébe is. Emellett felvettem a kapcsolatot közép-európai partnerintézményekkel későbbi intenzívebb együttműködési formák megalapozása céljából. Önálló fejezetben tekintem a dolgozatban át a nemzetközi érintettségű kutatómunka keresztmetszetét. A válogatás szempontja Európa-központúságot tükröz, elsősorban abból a szempontból válogattam, hogy honnan lehet a leginkább jó gyakorlatokat átvenni a magyar projekt számára. A nemzetközi háttér kutatása során az amerikai és az európai webarchiválás modelljei között jelentős különbségeket tapasztaltam. Európában alapvetően nemzeti közgyűjteményekre hárultak a webarchiválás feladatai, néhány egyetem, kutatóintézet csatlakozásával. Az Egyesült Államokban a legtöbb intézmény, legyen az közgyűjtemény, egyetem vagy kutatóintézet csupán webkurátori tevékenységet végez, a webarchiválási feladatok technikai lebonyolítását átadják az Internet Archive kereskedelmi üzletágaként működő Archive-IT nevű vállalkozásnak (Leach-Murray, 2018). A nagyobb egyetemek, közgyűjtemények, kutatóintézetek pedig maguk építik ki digitális repozitóriumi rendszereiket, s ebbe integrálják bele a webes állomány megőrzését is. Az IIPC rendezvényeken sokszor hallottam olyan véleményeket amerikai kutatók, webkurátorok részéről, hogy nagyon hálásak, hogy a webarchiválás technikai feladatai lekerülnek a vállalkról. Másrészt viszont azért fájlalják, hogy nincsenek olyan szakemberek a soraikban, akik az európai és egyéb kontinensekről érkező kollégák partnereiként technikai kérdésekben, a fejlesztési irányvonalak formálásában kompetens módon meg tudnának nyilvánulni. Mindenesetre lényeges különbség a könyvtári webarchiválási gyakorlatot tekintve, hogy a technikai háttérrel kiszervezik, s csupán a gyűjteményfejlesztés, szervezés marad az intézmény hatásköre, vagy pedig a teljes infrastruktúrát is egy adott intézmény bázisán üzemeltetik. Le kell szögezni, hogy a nemzeti könyvtáraknak szinte minden vizsgált esetben központi szerepköre van a webarchiválási szolgáltatások szervezésében. E tevékenységek szervezeti keretei azonban rendkívül változatosak. A siker biztos garanciájának a tartalomszolgáltatók és közgyűjtemények közötti széleskörű összefogás tűnik. A webarchiválás közös ügy, nem egyetlen intézmény felelősségi körébe tartozó szolgáltatási feladat. Az adott ország közigazgatási szerkezete, a

közművelődési intézmény struktúra jellege alapvetően meghatározza a webarchiválási tevékenység szervezeti kereteit (Németh, 2017). A webtechnológia robbanásszerű, folyamatos fejlődésével egyre nagyobb kihívást jelent a különféle webes tartalomszolgáltatások archiválhatóságának biztosítása, mind a tartalomszolgáltatók, mind az archiváló intézmények részére. Ígéretes fejlődési irányt jelent, hogy a tartalmakat egyre nehezebben begyűjteni képes aratórobotok helyett a böngészőprogramok működését imitáló, a böngészőmotoron alapuló új szoftverek jelennek meg, ahol már a felhasználó képernyőjén megjelenő tartalom rögzítése válik céllá. Ilyen például a Brozler nevű szoftver, melyet a Francia Nemzeti Könyvtár munkatársa mutat be tanulmányában. (Sylvie-Rollason-Cass, 2020). Ez részleges paradigmaváltást is hozhat a jövőben, a webarchiválás területén alapvetően átalakítva a jelenlegi nemzeti modellek gyakorlati kereteit.

3.2 A nemzetközi kutatás-fejlesztési eredmények implementálása a magyar webarchiválás gyakorlati tevékenységi keretei közé

A legújabb technológiák, szoftverek, gyakorlatok átvételével jelentősen mérsékelni lehetett a kései magyarországi kezdésből adódó hátrányokat. A negatív hatások inkább azon a téren jelentkeznek, hogy a magyar web 2017-ig terjedő történetének tanulmányozás során teljesen az amerikai Internet Archive által megőrzött anyagokra vagyunk utalva. Remélhetőleg ezt az anyagot hamarosan sikerül megvásárolni és/vagy szolgáltatathatóvá tenni egyedi, az OSZK logójával ellátott kezelőfelületen keresztül. A bevezetésben, s az egyes közgyűjteményi ágakkal foglalkozó fejezetekben széleskörű áttekintést nyújtok a webarchiválással kapcsolatos elméleti és gyakorlati kihívásokról.

A közgyűjtemények szempontjából fontos állításom, hogy a webarchiválás semmiféleképpen nem jelenik meg radikálisan új feladatként. A digitális könyvtári, levéltári, múzeumi munkafolyamatokat kiterjesztő, azokhoz szervesen illeszkedő tevékenységről van szó. Meglátásom szerint, mivel a kulturális örökség különféle területei új felületen jelennek meg, voltaképpen a hagyományos gyűjtőköri kereteket kell kiterjeszteni a webarchiválás irányában. Önálló fejezet ad áttekintést ezekről a kiterjesztési lehetőségekről, a webarchiválás könyvtári, levéltári múzeumi felfogásáról, feladatairól. Alapvető fontosságú a tézisekben foglaltak szempontjából is a webarchiválás teljes közgyűjteményi környezetének

bemutatása, és annak vizsgálata, hogy milyen értelmezésekben jelenik meg a webarchiválás feladatként az egyes közgyűjteményi ágak szintjén.

A nemzetközi kutatás-fejlesztési eredmények hasznosításában is jelentős eredményekről tudok számot adni. Az archivált állomány teljesszövegű keresését biztosító SolrWayback szoftver fejlesztése magyar-dán együttműködésben zajlik, a magyar webarchívum adja annak nyilvánosan bemutatható szolgáltatási felületét a dániai szolgáltatások jogi korlátai miatt. Az International Internet Preservation Consortium (IIPC) tagjaként az OSZK-ban is első kézből értesülünk a webarchiválással kapcsolatos nemzetközi kutatás-fejlesztési eredményekről. A korszerű webarchiválási hardveres és szoftveres háttér kialakítása s jelenlegi szolgáltatási palettája önálló fejezetben kerül megtárgyalásra a dolgozatban. A webarchiválási szolgáltatás kereteit adó kísérleti projekt megalapozásához nagy segítséget nyújtott az az online kurzus, melyet az Aarhushi Egyetem NETLAB kutatócsoportja keretébe végeztem el 2017-től. A későbbiekben sikerült a Web Curator Tool nevű webarchiválási munkafolyamat kezelő rendszer tesztelésébe és használatába is bekapcsolódnai, jó személyes kapcsolatokat kialakítva a fejlesztőkkel. A dolgozatban részletesen számot adok azokról az archiválási munkafolyamatokról is, melyeket a nemzetközi minták szerint alakítottunk ki a nemzeti webtér általános, pillanatfelvételszerű archiválására, illetve az adott témakör vagy esemény kapcsán végzett válogatott tartalmakra kiterjedő webarchiválási tevékenységre.

Végeredményben elmondható, hogy rövid idő alatt sikerült sokszínű nemzetközi kapcsolatokat kiépíteni, s ezek szerepe az állandó szakmai tapasztalatsere lehetőségével (a nemzetközi szakirodalom áttekintése mellett) kritikus fontosságú volt a magyarországi webarchiválás szakmai alapjainak meghatározásakor. Fontos tehát a dolgozatban dokumentálni mind a nemzetközi kapcsolatok kialakításának lépéseit, mind annak eredményeit is az informatikai háttér és a munkafolyamatok vonatkozásában.

Kapcsolódó előadások, publikációk

- Németh, M. Nemzetközi körkép a webarchiválás gyakorlatáról. Könyvtári Figyelő, 63(4), 575–582, 2017
- Németh, M. A webarchiválás nemzetközi környezete: Mozaikok az IIPC 2019 kongresszusról. Könyv, Könyvtár, Könyvtáros, 27(12), 23–27, 2017

- Németh, M. Hogyan tudjuk fejleszteni a webes gyűjteményünket? A Holland Nemzeti Könyvtár webarchiválási tevékenységének értékelése (2007–2017). Könyvtári Figyelő, 64(4), 634, 2018
- Németh M. Webarchiválás két szakmai rendezvény tükrében. Könyv, Könyvtár, Könyvtáros 28(6), 26-29, 2019
- Geeraert F. & Németh M. Exploring special web archives collections related to COVID-19: The case of the National Széchényi Library in Hungary, WARCnet Papers, Aarhus, 2020.
- Németh M. A short introduction to the Hungarian Web Archiving Pilot Project. IIPC General Assembly, Zagreb, 2019
- Drótos, L., Németh, M. (2019.06.07). Metadata supported full text search in a web archive. IIPC Web Archiving Congress, Zagreb, 2019
- Németh, M. The first two years of the Hungarian Web Archiving Pilot Project. Colloquium of Library and Information Experts of the V4+ Countries, Bratislava, 2019

[T4] A webarchiválás tágabb összefüggésrendszerének ábrázolása, az archivált anyagokra épülő kutatások különféle útjainak bemutatása, a webarchiválás itthoni korszerű gyakorlati keretei megteremtésének felvázolásával együtt, összetett tudományos igényű elemzés elvégzését teszi lehetővé. Önálló tudományos tevékenység keretében fontos felvázolni a jövőbe mutató kutatási irányokat is.

4.1 A webarchívum mint a tudományos kutatások tárgya

A tágabb összefüggésrendszer részeként a webarchiválás eredendő töredékességéről, a technikai kihívásokról, az alapvető információtudományi, közgyűjteményi összefüggések felvázolásáról már szó esett korábban. A webes kulturális örökség archiválásának védelmére, fontosságára, teendőire, a közgyűjteményi tevékenységek fontosságára több átfogó nemzetközi dokumentum, tanulmány is rámutat, a dolgozat keretei között kiemelten ismertetem az UNESCO 2003-ban megfogalmazott chartájában foglaltakat (UNESCO, 2003). Ez leszögezi, hogy „A világ digitális örökségét a végleges eltűnés veszélye fenyegeti. Ehhez hozzájárul az azt tartalmazó hardver és szoftver gyors elavulása, bizonytalanság a forrásokkal, felelőségekkel, és a fenntartás és megőrzés módszereivel kapcsolatban, valamint a támogató jogi szabályozás hiánya.” A közgyűjtemények felelőségéről pedig így szól: „A nemzeti megőrzési politikák kulcselemeként az archívumokkal kapcsolatos törvényhozásnak, és a

könyvtárak, levéltárak, múzeumok, és más nyilvános gyűjteményeknek fel kell ölelniük a digitális örökséget.”

A hagyományos megőrzési feladatokon túl egyre fontosabb bemutatni, hogy a webarchívum miként jelenhet meg a tudományos kutatások tárgyaként. Nagyon fontos felvillantani, hogy társadalmi szinten mi lehet az a hozzáadott érték, amely a webarchívum kutatása kapcsán a felszínre kerülhet. A közgyűjtemények szempontjából nagyon fontos új perspektívákat kínál az új kutatási irányok megalapozása, külső partnerségi formák feltárása. A tudományos és társadalmi presztízs emelkedését kínálja intézményi szinten is egy-egy jól megalapozott projektben történő közreműködés. A felsőoktatási intézmények számára pedig fontos új szinergiák tárulhatnak fel. A tudományos hasznosítással foglalkozó fejezetben először kitérek a kutatási célú felhasználás körül formálódó új WARCNET nemzetközi projektre, majd ezt követően egy szinte napjainkban született új tudományágot, a webtörténetírást mutatom be. A második témakörben a webarchívumról mint a digitális bölcsészeti kutatások tárgyáról teszek említést, a webarchívumban tárolt nagy mennyiségű adatkészletek tudományos kutatási célú felhasználásáról, illetve az archivált adatok vizuális megjelenítéséről esik szó adatbányászati és adatelemzési megközelítésben. A webarchiválás, illetve az archivált tartalom visszakeresése hatékonyságának növelésére vonatkozó kutatási tervek felvázolása szemantikus webes eszközök segítségével a hatodik fejezetben kap helyet. Ez a mikroadatokkal foglalkozó fejezet olyan további kutatásokat segítő témafelvetéseket vezet elő, melyek hosszabb távon talán az egész dolgozat legfontosabb hozzáadott értékkel bíró tartalmi elemeiként jelenhetnek meg az információtudomány, illetve az informatika oktatásának-kutatásának kapcsán.

A webarchívumokra mint gyűjteményekre irányuló tudományos kutatásokat, illetve a webarchívumok gyűjteményeinek tudományos kutatási célú hasznosítását európai szinten a WARCnet projekt keretében fogják össze. Az Országos Széchényi Könyvtár (OSZK) részéről 2020 őszének végén, a második online konferenciát követően tudtunk csatlakozni a projekthez. Résztvevői európai nemzeti könyvtárak webarchiválást végző munkatársaiból, tudományos kutatóiból, egyetemi oktatást-kutatást végző személyekből (elsősorban kommunikációkutatók és történészek), illetve informatikus fejlesztők közül kerülnek ki. Négy munkacsoport keretében folyik a tevékenység. Az első

munkacsoport a különböző intézményi webarchívumok gyűjteményeit érintő összehasonlító kutatásokra fókuszál. A második munkacsoport a több webarchívum által közösen épített nemzetközi gyűjteményekben rejlő kutatási, elemzési lehetőségeket tárja fel. A harmadik munkacsoport a webarchiváláshoz szükséges információtechnológiai fejlesztések területén mozog, a negyedik pedig a nyílt adatok menedzsmentjének webarchiválási vonatkozásait tartja szem előtt, beleértve a téma tágabb digitális közgyűjteményi vonatkozásait is. Például a webarchívumok integrációját integrált könyvtári rendszerekbe, teljesszövegű keresési funkciók fejlesztését. (Németh, 2020).

A webtörténetírás egy nagyon fiatal, igazából a 2010-es években kibontakozó tudományág. Önálló tudományos folyóirattal is rendelkezik már *Internet Histories* címmel, első számának bemutatkozó tanulmánya átfogó bemutatással szolgál e tudományos területről Niels Brügger és nemzetközi kutatócsoportjának segítségével (Brügger, Goggin, Milligan, & Schafer, 2017). A kutatások tárgya nagyon széleskörű témaköröket foglal magában. A web múltjának tanulmányozása kulcsfontosságú a jelen fejlődési tendenciáinak értelmezéséhez is. Ez nem csupán a szűken vett világháló múltjára érvényes, hiszen az internet fejlődésének politikai, gazdasági, társadalomtörténeti aspektusai is kulcsfontosságúak a jelen trendjeinek elemzése szempontjából (Brügger, 2013). A történészeknek ugyanúgy fontos szerepet kellene játszaniuk a webarchiválás kutatási célú felhasználása mögött álló archiválási intézményrendszer kereteinek meghatározásakor, mint ahogyan az a hagyományos levéltárak szervezeti rendje és munkafolyamatai esetében a 19. században történt. Fontos feladata a történészeknek a webarchívumok biztonságának vizsgálata is abból a nézőpontból, hogy a múlt archivált tényeinek manipulálását, aktuális politikai célokra történő újírását is meg kellene akadályozni. Itt fonódik össze egymással az IT-biztonság és a történelmi hitelesség szempontrendszer. A számítógépes világhálónak mint technikai infrastruktúrának, illetve mint kommunikációs és publikációs platform történetének tanulmányozása is hangsúlyosan előtérbe kerül. Emellett egy adott személy, esemény, témakör, intézmény webes lenyomatának nyomon követésére is lehetőség nyílik. A webes információ nagyon gyorsan avul, a digitális műveltség, illetve a személyes információk kezelésének fontos eleme lenne a személyes archiválás módszereinek széles körű oktatása és alkalmazása.

A Networkshop 2020 című digitális könyvtári és informatikai konferencián, 2020 őszén önálló tutoriált szerveztem a témakör bemutatására. Emellett a középiskolás korosztálynak digitális tananyagot is fejlesztünk munkatársaimmal együtt e témakörben.

A fentebb tárgyalt témaköröket egészíti ki az archivált szöveges, illetve vizuális webes tartalmak, vagy akár adott webszerverek naplófájljainak tanulmányozása a gépi tanulás (machine learning) eszköztárával, illetve a nagymennyiségű adatok elemzésének módszereivel (Brügger, 2009). A kutatás szintjeként megjelenhet egy egyedi webes fájl vagy weboldal, illetve egy adott webhely, doméntartomány is. Legtágabb értelemben pedig a webes univerzumnak, mint olyannak a történetét is vizsgálni lehet. A Memento protokoll segítségével több webarchívum archivált anyagai is összevethetővé válnak egy adott weboldalról, illetve témakörről, illetve ennek szükségszerű korlátait is feltárták már (Brunelle, Kelly, Salaheldeen, Weigle, & Nelson, 2015; Klein, Shankar, Balakireva, & Van de Sompel, 2019). Egy másik nézőpont az, amikor azt vizsgáljuk, hogy ki és milyen céllal, hatókörrel folytat webtörténeti kutatásokat. Amikor kis mennyiségű forrásanyagot, akár csupán egy meghatározott honlap történetét tanulmányozzuk egy adott szoftverkörnyezettel, speciális kutatási céllal, azt nevezzük Niels Brügger terminológiájával élve mikroarchiválásnak.

Fontos megemlítenem ebben az összefüggésben azt is, hogy miután maga a webarchiválás eredendő módon egy összefüggő egészből csak töredékeket tud kiragadni, ez óhatatlanul hatással van a webtörténészek tevékenységére is. Azzal kell dolgozniuk, ami a rendelkezésükre áll, s adott esetben megpróbálni pótolni a múlt hiányzó darabkáit. Ebben persze semmiféle nívum sincs, hiszen a hagyományos történeti források használata is általában hasonló dilemmákat vet fel. Ami mégis különlegessé teszi a webtörténetírói munkát a hagyományos történetírói tevékenységhez képest, hogy a vizsgálódásaink módszertani hátterének megtervezésének van egy speciális összetevője. Egyaránt tisztában kell lenni a webarchívumban tárolt források jellegzetességeivel, a webarchiválás munkafolyamatának főbb jellemzőivel, illetve annak a hardver- és szoftverkörnyezetnek a sajátosságaival, illetve korlátaival, melyek révén vizsgálódásainkat folytatni tudjuk.

A webtörténészek munkájának támogatására egyre inkább előtérbe kerül speciális munkakörnyezetek kialakítása. Egyre súlyosabb kihívásként jelent meg, hogy a történettudósok közül sokan nem szándékoznak mélyreható informatikai ismeretekre szert tenni, miközben számukra is biztosítani kellene a webarchívumokban tárolt anyagok kutatási célú elemzését. A történészek webarchívumokat érintő kutatásait segítő projektről részletesebb ismertető is készült (Németh, 2020c), itt csak egy szűkebb áttekintésre szorítkozunk. A könyvtárak, levéltárak, múzeumok digitális gyűjteményeinek használatához a GLAM Workbench szolgáltatási környezet eddig is biztosított eszközöket, gyakorlati témákat és leírásokat. Ehhez kellett hozzáilleszteni a webarchiválás témakörét (Németh, 2020c). A cél, tehát annak bemutatása lett, hogy a történeti kutatások során felmerülő témákat a webarchívumokból nyert adatok elemzésének segítségével miként lehet újszerű nézőpontokból feltárni („Jupyter notebooks for web archives”, 2020). Ehhez úgynevezett Jupyter digitális munkafüzeteket (notebook) hoztak létre. Az elméleti háttérrel tartalmazó szöveges útmutatók mellett részletes leírások is találhatóak gyakorlati példákkal a különféle elemzési módszerek használatáról („Asking questions with web archives – introductory notebooks for historians—IIPC”, 2020). Az igazán újszerű megoldást azonban egy webböngészőbe integrált gyakorlófelület jelenti, melynek segítségével hozzáférhetünk egy adott webarchívumhoz és a gyakorlatban is kipróbálhatjuk, hogy miként lehet adatbányászati, illetve adatelemzési tevékenységeket folytatni, konkrét példákon keresztül lehet körbejárni különféle felhasználási lehetőségeket. Sőt arra is van lehetőség, hogy ezeken a gyakorlófelületeken különféle előregyártott alkalmazásokat futtassunk le, s mérjük fel azok kimenetének felhasználási lehetőségeit („Jupyter notebooks for web archives”, 2020).

A webarchívumok nagy szövegtömegű tárházaként adattudományi projektek középpontjában is állhatnak. Az értékes adatok gyors feldolgozása, illetve visszakeresésének biztosítása egyre inkább előkerül a webarchívumok használata során. Ilyen adatok lehetnek a naplófájlok adatai, speciális tranzakciós (pl. geológiai adatok) vagy különféle az adott archív gyűjteményben tárolt szöveghez kötődő adattípusok is (Lnenicka, Hovad, & Komarkova, 2015). A webarchívumok a disztributív adatfeldolgozáshoz is segítséget nyújthatnak

Apache Hadoop segítségével egy megadott alkalmazáskészlettel, adott platformon. A részlegesen strukturált, illetve teljesen strukturálatlan adatkészletek számos kutatási célból vizsgálhatók. Fókuszálhatunk a webes tartalmakra, a tartalomhasználati adatok kinyerésére, illetve a webes szerkezeti elemek feltárására is. A tartalomfeltárás egyes webhelyek, illetve webhelycsoportok által közölt információk visszakeresésében segíthet. A fő cél az, hogy strukturált adatokat nyerjünk ki ezekből a tartalmi erőforrásokból. Ezek az adatforrások aztán integrálhatók szemantikailag hasonló adatelemekkel, valamifajta tartalmi hierarchia vagy tartalomintegráció alkotható meg a segítségükkel. (Lnenicka és mtsai., 2015). A különféle strukturált, illetve részben strukturált adatkészletek kvantitatív alapú történeti elemzések tárgyául is szolgálhatnak. Ebben az esetben az adattudós kinyeri az adatokat a webarchívum gyűjteményéből, s segítséget nyújt a webtörténésznek azok elemzésében is.

Egy újabb érdekes alkalmazási területet villant fel a 2020 végén zárult LinkGate nevű projekt, mely a webarchívumokban tárolt nagy mennyiségű adatok vizualizációjával foglalkozik. A projekt három alap komponensből áll. Az első egy Link-indexer névre hallgató indexelő eszköz („IPC RSS LinkGate Webinar”, 2020). Ez kinyeri a szükséges metaadatokat a webarchívumokban tárolt WARC fájlokból (WARC fájl URI címe, WARC fájl dátuma, az adott webhelyről kifelé mutató linkek listája), s egy önálló WAT névre hallgató fájltypusban tárolja azokat. A Link-serv komponens a Link-indexer által kinyert adatokat szemantikus adattárban (data store) tárolja el, s gráf alapú adatsémát rendel hozzá (Németh, 2020c). A gráf alapú adatbázist a Neo4j nevű No-SQL adatbáziskezelő rendszer menedzseli („IPC RSS LinkGate Webinar”, 2020). A Link-Viz nevű harmadik komponens pedig egy megjelenítési felületet biztosít, ahol az adatbázisban tárolt adatok webböngészőn keresztül vizuálisan gráf-adatszerkezetben megjeleníthetők („IPC RSS LinkGate Webinar”, 2020). Az egyes webhelyek egy-egy csomópontot alkotnak a hálózatban a közöttük lévő kapcsolatok pedig térben és időben is tanulmányozhatóvá válnak.

A nemzeti könyvtárak hatókörén általában kívül eső komponens a webarchívumokból történő hitelesített adatok szolgáltatása. Az archivált hiteles jogi dokumentumok gyűjtése és felhasználása az üzleti és a közigazgatási szféra szintjén jelenik meg. Számos országban (például Nagy-Britanniában vagy

Ausztráliában) törvény írja elő, a cégek teljes online tevékenységét (ideértve a közönségkapcsolati csatornák forgalmát a közösségi média felületekkel együtt) archiválni kell, a hiteles archivált anyagot jogi eljárásokban felhasználhatóvá kell tenni.

A dolgozat talán leginkább jövőbe mutató része a kutatások szempontjából a honlapok forráskódjába illeszthető szemantikus mikroadatokkal foglalkozik. Ennek a fejezetnek a legfőbb célja a további kutatómunka megalapozása a webarchiválás hatékonyságának növelése terén, olyan adatkészletek kidolgozásával melyek szabványos szemantikus jelölési tulajdonságokkal felvértezve beilleszthetők lennének bármely webhely forráskódjába és értelmezhetők lennének a webarchiválást végző szoftverek által. Elsőként definiálok a dolgozatban azt, hogy milyen típusú adatkészletekről beszélünk, majd a későbbiekben rátérek azok használatának jelentőségére a webarchiválás során.

A nemzetközi kutatási irányok, jövőbe mutató kezdeményezések felvázolása mellett a dolgozat hangsúlyos célja az eddig elért magyarországi eredmények összegzése is. A dolgozatban átfogóan felvázolom az eredeti szakmai koncepciót, kitérve az egyes részletek megvalósulásának módjaira. Szerepet kap a gyűjtőkör meghatározása, a webarchiválási munkafolyamatok bemutatása, a metaadatok rögzítési lehetőségeinek felvázolása, a minőség-ellenőrzés kérdései, s a hosszútávú megőréssel kapcsolatos teendők. A jogi háttér és a szabályozási keretek felvázolását követően nagyon fontos szót ejteni arról is, hogy az egész tevékenységi környezetet milyen kommunikációs keretrendszerben mutatjuk be a társadalom számára.

A kettős látásmóddal, az itthoni szolgáltatási keretek felmutatásával, valamint a nemzetközi kutatási célú felhasználás s a jövőbe mutató fejlesztési lehetőségek mintáinak felmutatásával bemutathatóvá vált tehát, hogy hol tartunk most, s milyen fejlesztési lehetőségek nyílhatnak a közeljövőben nem csupán a webarchiválással foglalkozó szakemberek, hanem más tudományágak képviselői számára is.

Kapcsolódó előadások, publikációk

- Németh, M. 404 Not Found – Ki őrzi meg az internetet; Webarchiválás workshop az Országos Széchényi Könyvtárban. Tudományos és Műszaki Tájékoztatás, 64(11), 577–582, 2017
- Németh, M. A webarchiválásról történeti megközelítésben. Könyv, Könyvtár, Könyvtáros, 27(2), 48–52, 2018
- Németh, M. Webarchívum mint a tudományos kutatások tárgya. Tudományos és Műszaki Tájékoztatás, 67(12), 757–765, 2020
- Drótos L. & Németh, M. Egyedi mentésekre szolgáló webarchiváló szoftverek. Könyv, Könyvtár, Könyvtáros 29(12), 3-11, 2020
- Németh, M., Drótos, L.: Web Archives as a Research Subject. In: Information and technology transforming lives: connection, interaction, innovation : Proceedings of the XXVII. BOBCATSS Symposium. Ed.: Gordana Gašo, Mirna Gilman Ranogajec, Jure Žilić, Madeleine Lundman, University of Osijek, Osijek, 471-478, 2019. ISBN: 9789533141213
- Németh, M. Webarchívum mint a tudományos kutatások tárgya. Networkshop 2019 konferencia, Győr, 2019.
- Németh M. Introduction to web archiving in digital humanities context. DH 2019 konferencia, ELTE, Budapest, 2019.

[T5] A webarchiváláshoz kötődő tevékenységek és elméleti kutatási irányok, keretek közgyűjteményi szempontú feldolgozására, a legújabb nemzetközi kutatási eredmények bemutatására, ötvözve a saját itthoni kutatói tapasztalatokkal, önálló tudományos munka formájában magyar nyelven korábban nem került sor. Új interdiszciplináris kutatási irány alapozható meg ezen hiányosságok pótlásával a magyar tudományos életben.

5.1 A webarchiválás megközelítésmódjai

A dolgozat készítése során sajátos kihívásként jelent meg, hogy a bemutatandó szolgáltatási gyakorlat nemzeti könyvtári jellegű, míg a webarchiválásnak mint tudományos témakörnek önálló könyvtári, levéltári és múzeumi vonatkozásai is vannak, emellett pedig a tudományos felhasználás kérdései is önálló bemutatást igényeltek.

A webarchiváláshoz kötődő gyakorlati tevékenységek ellátása mellett alapvető fontosságú a webarchiválás jelentőségének tudatosítása, a területet érintő oktatásmódszertani keretek felvázolása, a webarchiváláshoz kötődő

kompetenciák széleskörű terjesztése a közgyűteményi szakmai közeg és a tágabb társadalmi célközönség felé is. Az oktatási dimenzió, s annak itthoni vonatkozásai tehát önálló témakörben jelennek meg. Az oktatási terület fontosságának további indokaként felhozható még, hogy alkalmam nyílt aktívan bekapcsolódni a webarchiválás itthoni oktatási kereteinek formálása mellett a területet érintő nemzetközi kezdeményezésekbe is.

A webarchívumok egyre nagyobb fontosságra tesznek szert úgy is, mint a különféle tudományos kutatások forrásai. Ezzel az izgalmas területtel azért is érdemes foglalkozni, mert hihetetlenül széleskörű az az összefüggésrendszer, mellyel a webarchiválási tevékenységet és a webarchívumokat mint a kutatások forrását vizsgálni lehet. A webarchiválás gyakorlatának s a területhez kötődő oktatási háttérnek a felvázolása mellett tehát magas hozzáadott értékkel bíró eleme a disszertációnak, hogy a terület tágabb interdiszciplináris környezetéhez kötődő jelenkori és jövőbeni kutatási lehetőségeket, témákat mutasson fel. Egy konkrét javaslat is megfogalmazásra kerül a zárófejezetben, hogy miképpen lehetne a szemantikus technológiákban rejlő lehetőségeket a közgyűteményi szféra, illetve a tartalomipar összefogásával a webarchiválás terén hasznosítani. Így válik teljessé a webarchiválás környezetét, kapcsolódási pontjait felvázoló összkép. A dolgozat egyediségét annak közgyűteményi szempontú elemző látásmódja, valamint az itthoni intézményes webarchiválás kezdeteihez tartozó kutatás-fejlesztési munka eredményeinek bemutatása adhatja tehát meg a vonatkozó oktatási és kutatási-hasznosítási háttér alapjainak felvázolásával együtt.

A webarchiválás azért is tűnik rendkívül izgalmas kutatási területnek mert nagyon sokféle megközelítésmódot kínál. Tanulmányozhatjuk filozofikus megközelítéssel általában azt, hogy miként lehet egy önmagában szerves egységet alkotó organizmusból egyes elemeket önkényesen kiragadni. Miért nem lehet menteni akár az egész webet a maga teljességében? Miként lehet a reménytelenül összetett, ugyanakkor egyre több belső ökoszisztémára szétbomló webes világ egyes elemeit kiemelve a gyakorlati digitális megőrzésből adódó konkrét feladatokhoz hozzárendelni. Milyen módon lehet egyáltalán tanulmányozni az egyes részelemeket, részelemhalmazokat anélkül, hogy annak természetes környezetében létezik a webes univerzum részeként.

5.2 A webarchiválás közgyűjteményi környezete

Praktikusabb vizekre evezve a webarchiválás értelmezhető úgy is, mint digitális könyvtári, levéltári, muzeológiai feladat. A legfőbb cél a web egy adott, a kulturális örökség megőrzése érdekében fontosnak tartott szegmenseinek megőrzéséről gondoskodni. Mennyire illeszkedik ez a tevékenység a többi hagyományos, illetve digitális, könyvtári, levéltári, múzeumi munkafolyamathoz? A közgyűjtemények szempontjából fontos állításom, hogy a webarchiválás semmiféleképpen nem jelenik meg radikálisan új feladatként. A digitális könyvtári, levéltári, múzeumi munkafolyamatokat kiterjesztő, azokhoz szervesen illeszkedő tevékenységről van szó. Meglátásom szerint, mivel a kulturális örökség különféle területei új felületen jelennek meg, voltaképpen a hagyományos gyűjtőköri kereteket kell kiterjeszteni a webarchiválás irányában.

A könyvtárakban végzett webarchiválás gyűjtőköre alapvetően a világhálón nyilvánosan publikált webes tartalomra terjed ki. A webarchívum mint gyűjtemény mozaikszerűen illeszkedik be a digitális könyvtári szolgáltatások környezetébe.



3. ábra A webarchiválás helye a digitális könyvtári szolgáltatások rendszerében

A saját nemzeti könyvtári webarchívum tervezésekor, a nemzetközi gyakorlatból is adódó tanulságokat leszűrve, hamar tudatosítanom kellett magamban, hogy alapvetően kétféle válogatási elvről beszélhetünk. A tematikus gyűjtemények összeállításáról (ez lehet egy adott téma, illetve esemény, történes, rendezvény is), illetve pillanatfelvételek készítéséről a webnek a .hu domain alatti (illetve egyéb

legfelső szintű domaineken található magyar nyelvű) anyagairól, ami tematikus aratásnál jóval felületesebb, áttekinthető jelleggel valósul meg.

A levéltári webarchiválás tekinthető véleményem szerint a szó szoros értelmében vett webarchiválási tevékenységnek, bár kiterjedtsége, nagyságrendje jellemzően elmarad a könyvtárakban végzett webarchiválási tevékenység mögött (Drótos & Németh, 2018b). A könyvtári és a webarchiválási tevékenységek közötti legélesebb választóvonalat a gyűjtőkörben és az archiválási módszerekben lehetjük fel, mint ahogy hagyományosan is ezek a jellemzők különítik el a két intézménytípust egymástól. A könyvtári területhez hasonlóan itt is a hagyományos tevékenységek új szempontok szerint történő kiterjesztéséről van tehát szó a web, mint információhordozó médium megjelenésével. A levéltárak hagyományos szerepkörükből adódóan elsősorban a nem a széles nyilvánosságnak szánt publikálási szándékkal létrehozott webes tartalmakat mentik. Az intézményi kommunikáció fontos színtereivé váltak például a fórumok, személyes blogok, e-mailek, fotók, videók, személyes webkettes oldalak és csatornák, csevegőprogramok forgalma, zárt csoportok és webhelyek. Ám a hagyományos levéltári tevékenységhez talán a leginkább illeszkedő gyűjtőköri elem az intézményes és céges belső honlapok és iratok kezelése. Ezt a gyűjtőmunkát igazából el sem lehet választani az adott szervezetre vonatkozó általános levéltári gyűjtési tevékenységektől, azoknak egy szeleteként jelenik meg. A begyűjtött anyagot a levéltári feldolgozási rend szerint digitális fondokba szervezik, az adott szervezetre, személyre, témakörre vonatkozó többi digitális tartalommal együtt. A webarchiválás múzeumi szegmense gyűjtőkörét tekintve erőteljesen eltér a könyvtári, illetve levéltári szegmenstől. A középpontban itt a művészeti, illetve iparművészeti tevékenység webes lenyomatainak megőrzése, illetve megjeleníthető állapotban tartása áll (Drótos & Németh, 2018b). A web, mint utaltam arra korábban már, önmagában rendkívül sérülékeny, s fokozottan igaz ez a webes művészeti alkotásokra. Ezek általában szorosan kötődnek egy adott korszak internetes technológiai elemeihez, a szolgáltatási környezet átalakulásával, elavulásával könnyen elérhetetlenné válhatnak. Épp emiatt az elavuló környezet miatt a megtekinthető állapotban tartás, a megfelelő hardveres, szoftveres emulációs lehetőségek megtalálása is komplex feladatokat állít elénk. (Drótos & Németh, 2018b). A múzeumi webarchiválás két fő szegmensét erre

tekintettel véleményem szerint az alábbi módokon határozhatjuk meg. Az első az egyes weben alkotó művészek életművének minél szélesebb körű felkutatása és archiválása. A második pedig az adott múzeum gyűjtőkörének megfelelő szelektív webarchiválási tevékenység. Ennek határai rendkívül szélesek, s igazodnak a múzeumi intézményi terület színes palettájához. (Drótos & Németh, 2018b).

5.3 A webarchiválás technikai eszköztára

A webarchiválásra tekinthetünk úgy is, mint olyan tevékenységre melyhez nélkülözhetetlen informatikai eszköztár tartozik mind hardveres mind szoftveres vonatkozásban. Át kell tehát tekinteni mind annak a hardver infrastruktúrájának az alapvető elemeit, amely az archivált állomány tárolását, elérhetőségét biztosítja. Másfelől a webarchiválási tevékenység, illetve az archivált állomány visszanezhetőségének nélkülözhetetlen eleme az a szoftveres háttér, melynek segítségével az egyes tevékenységeket voltaképp el lehet végezni. A két fő szoftveres kategória tehát az archiváló és a megjelenítő szoftverekhez kötődik. A nemzeti könyvtárakban működő webarchívumok esetében általános módon az OSZK részéről is alapvetően olyan nyílt forráskódú szoftverekből tervezzük felépíteni a rendszert, amiket az Internet Archive és az IIPC nevű nemzetközi konzorcium is támogat. Ezeket fogjuk kiegészíteni saját fejlesztésű modulokkal vagy önálló alkalmazásokkal, melyek speciális igényeket elégítenek ki, illetve integrálják a webarchívumot az Országos Könyvtári Rendszer keretében bevezetésre kerülő új könyvtári rendszerrel, a metaadat-nyilvántartással, a digitális raktárral, a névterekkel, a statisztika-készítő, a felhasználói azonosítást végző, a jogosultságokat kezelő, a dokumentum-hitelesítő, a partneryilvántartó és még további szóba kerülhető rendszerekkel.

5.4 A webarchiválás oktatása

Az internet archiválás felfogható úgy is, mint oktatási feladat. Ennek kapcsán a dolgozatban elsőként azokkal a nemzetközi mintákkal, együttműködési formákkal foglalkozom, amelyekből ötleteket tudunk meríteni a saját webarchiválási tevékenység megalapozásához. Ezt követően rátérek a Könyvtári Intézettel, illetve a webarchiválási projekttel párhuzamosan futó E-learning projekt munkatársaival együtt kidolgozott tanfolyami keretek felvázolásához. A közgyűjteményi szakemberek továbbképzésének biztosítása mellett fontos

jövőbeni célként jelenik meg a közoktatás szereplői felé történő nyitás, illetve a webarchiválás iránti széles társadalmi érdeklődés felkeltése. A jövőre utaló tervek között tehát egy olyan formálódás alatt lévő tananyagról, illetve képzési keretrendszerrel ejtek szót, amely a bárki által végezhető személyes webarchiválás kompetenciáinak elsajátításában, a kapcsolódó szoftveres eszköztár megismerésében és kezelésében segít.

5.5 Kutatástámogatás

Az internet archiválásról, mint a tudományos kutatások tárgyáról pedig már az előző tézis kapcsán részletesebben szót ejtettem webtörténeti, adatbányászati, adatvizualizációs megközelítésben. Reményeim szerint a Jupyter Notebook projektet hozzá tudjuk majd illeszteni az OSZK webarchívumához, s ezzel a nem kifejezetten informatikai vénájú történész, digitális bölcsész kutatók számára is új lehetőségek nyílnának meg a webarchiváláshoz kötődő kutatások kapcsán. A fejlesztők szándéka szerint a Memento protokoll, valamint a webarchívumok által használt, a dolgozat technikai háttérrel foglalkozó fejezetében ismertetett szoftverek (Heritrix, Brozler, OpenWayback, PyWB) használatával bármelyik nemzeti könyvtári webarchívum összekapcsolható a tananyagokkal („Jupyter notebooks for web archives”, 2020). Így akár arra is lesz lehetőség, hogy össze lehessen hasonlítani a különféle webarchívumok archiválási módszereit, a metaadatmodelleket, illetve az azokra épülő szolgáltatásokat. Lehetőség nyílik az időbeli dimenzió tanulmányozására, hogy miként fejlődött egy-egy webhely, vagy egy-egy témakör mikor jelent meg a weben s hogyan gyarapodtak az ahhoz kötődő honlapok számszerűleg, illetve tartalmilag egyaránt („Jupyter notebooks for web archives”, 2020). A fejlesztők szándékai szerint ez az oktatási környezet a zárt hozzáférésű archív gyűjteményekhez is hozzáilleszthető lesz. A kutatások támogatásában tehát reményeim szerint az OSZK is segítséget tud majd nyújtani, s egyfajta új közös platform nyílik meg az adattudósok és történészek igazán interdiszciplináris jellegű együttműködéséhez is.

Kapcsolódó előadások, publikációk

- Németh, M. Webarchívum mint a tudományos kutatások tárgya. Tudományos és Műszaki Tájékoztatás, 67(12), 757–765, 2020
- Drótos L. & Németh, M. Egyedi mentésekre szolgáló webarchiváló szoftverek. Könyv, Könyvtár, Könyvtáros 29(12), 3-11, 2020

- Németh, M., Drótos, L.: Web Archives as a Research Subject. In: Information and technology transforming lives: connection, interaction, innovation : Proceedings of the XXVII. BOBCATSS Symposium. Ed.: Gordana Gašo, Mirna Gilman Ranogajec, Jure Žilić, Madeleine Lundman, University of Osijek, Osijek, 471-478, 2019. ISBN: 9789533141213
- Drótos, L., & Németh, M. A blended learning-based curriculum on Web archiving in the National Széchényi Library. Digital Library Perspectives 35(2), 97-114, 2019
- Drótos, L. & Németh, M (2018). The education of web archiving. In Katarina Tomková (Eds.) CDA 2018 Trvalá udržateľnosť a perspektívy ďalšieho rozvoja LTP archívov. (pp.108-116). Bratislava: Univerzitná knižnica v Bratislave, 2018.

[T6] A webarchiválás mint tudományos témakör a magyar tudományos életben történő meghonosításának kritikus előfeltétele az egész kutatási területet átfogóan bemutató hiánypótló tudományos munka elkészítése, az itthoni és nemzetközi tudományos és gyakorlati jellegű erőfeszítések összefoglaló jellegű, elemző bemutatásával együtt. A közgyűjteményi szempontú elemző látásmód, az itthoni intézményes webarchiválás kezdeteihez tartozó kutatás-fejlesztési munka eredményeinek bemutatása, a vonatkozó oktatási és kutatási-hasznosítási háttér alapjainak felvázolásával együtt adhat lendületet az egyes webarchiválást érintő részterületek további kutatásához.

6.1 A webarchiválás közgyűjteményi háttére kutatásának tanulságai

A dolgozat egyik fő célja a webarchiválásról szóló általános hiánypótló jellegű áttekintés elkészítése. A webarchiválás egyes közgyűjteményi ágakat érintő fő vonatkozásainak áttekintését senki sem végezte még el magyar nyelven. Ennek néhány tanulságát összegzem az alábbiakban.

Már említettem korábban, hogy a könyvtárakban végzett webarchiválás gyűjtőköre alapvetően a világhálón nyilvánosan publikált webes tartalomra terjed ki.

A kötelempéldány szabályok kiterjesztésével egyre inkább lehetővé válik a webnek az egyes országokhoz, nemzeti kultúrákhoz kötődő szeletei archiválásának előírása is. Erre több nemzeti webarchívum (brit, dán, svájci)

munkatársai is utalnak, Michaela Mayr és Andreas Predikaka pedig osztrák részről kifejezetten hangsúlyozza a nemzeti szempontok érvényesülését (Bingham, 2014; Mayr & Predikaka, 2016; Schostag & Fonss-Jorgensen, 2012; Signori, 2017). Meglátásom szerint a hagyományos passzív befogadói szerephez képest a könyvtáraknak sokkal proaktívabb szerepet kell játszaniuk a gyűjtőkörükbe eső webes (sőt tágabb értelemben a digitális) tartalmak begyűjtésében. Miután bárki előállíthat online tartalmakat, aki nincs is tudatában annak, hogy rá is vonatkozhat kötelezpéldány rendelkezés, ezért egy jól szervezett könyvtári hálózat, a nemzeti könyvtárral az élén sokat tehet a gyűjtőkörileg leginkább releváns témakörök kapcsán a tartalmak feltérképezéséért és begyűjtéséért. Már hosszú évek óta nemzetközi szinten evidenciaként fogalmazódik meg, hogy azt az elképzelést, ami a hagyományos kötelezpéldány rendelkezések mögött húzódik meg, hogy a teljesség igényével szolgáltatassák be a könyv és sajtótermékeket, a web esetében fel kell adni (Schostag & Fonss-Jorgensen, 2012; Stirling és mtsai., 2012). Fogalmunk sincs róla, miből áll voltaképpen a teljesség a világhálón, valamint eszközünk sincs rá, hogy megragadjuk azt. Fontos tanulság, amivel a webarchiválás itthoni gyakorlatának kialakítása kapcsán is szembesültem, hogy egy könyvtár csak arról a webes információforrásról tud értesülni, ami a látókörébe kerül, vagy amiről tájékoztatják.

A kutatás egyik fontos tanulsága, hogy némi párhuzamosság is megfigyelhető a könyvtári és levéltári webarchiválási tevékenységek között a kormányzati és államigazgatási honlapok mentésével kapcsolatban. Nincs általános nemzetközi gyakorlat arra, hogy ezt az anyagot könyvtári (pl. parlamenti könyvtári) vagy levéltári (pl. nemzeti levéltári) keretek között, illetve könyvtári vagy levéltári módszerek szerint dolgozzák- e fel, s hogy bekerül-e az anyag a kötelezpéldány szolgáltatás körébe (Espley, Carpentier, Pop, & Medjkoune, 2014; „Government Web& Data Archive”, 2017; Grimshaw, 2016; Holub & Rudomino, 2015; Niu, 2012). Számos országban a digitális állományok archiválására és megőrzésére a teljes államigazgatás körét átfogó kormányzati stratégia készül (Grimshaw, 2016; Paterson, 2015; Storrar & Seaman, 2016). Ez magában foglalja a gyűjtőkör pontos szabályozását, az egyes intézmények általános teendőit, illetve az archiválásért felelős intézmények tevékenységének alapszabályait. A későbbi történeti

kutatások számára alapvető fontosságú, hogy az állam, a közigazgatás, az államigazgatás szervei és intézményei miképpen jelentek meg a virtuális térben, milyen online iratokat keletkeztettek. Ahol a levéltárak webarchiválást végeznek, ez a fajta megőrzési munka áll a digitális webarchiválási tevékenység fő középpontjában, a gyűjtőkör elemei csupán kiegészítő jellegűek. Ebben sincs nagy újdonság, a hagyományos levéltárosi tevékenységek prioritásai is e szerint alakulnak. Például kínálkozik a megőrzés fontosságára az az önálló kormányzati webarchívum szolgáltatás, amelyet a Brit Nemzeti Levéltár üzemeltet 1996 óta (National Archives, 1996). A szolgáltatás így szinte egyidős a legáltalánosabb jellegű webarchiválás tevékenységet végző amerikai Internet Archive-val. Ennek a webarchívumnak hatóköre nem csak a szokásos fent említett államigazgatási webes dokumentumkörre terjed ki, hanem mentik a különféle intézmények Twitteren, Flickr-en, egyéb közösségi média és videomegosztó csatornákon közzé tett információit is. Ám mint említettem egyáltalán nem törvényszerű, hogy az államigazgatás online anyagait levéltári keretek között archiválják, megjelenhet az a könyvtári gyűjtőkör részeként is. Arra is van példa, hogy széleskörű együttműködéssel kampányszerűen egy-egy elnöki terminus lezárulta körül gyűjtik egy-egy időszakra vonatkozóan a kormányzati weboldalakat, ez megfigyelhető például az USA-ban a Government Web & Data Archive project kapcsán („Government Web& Data Archive”, 2017; Phillips & Phillips, 2018). Sajnos Magyarországon levéltári webarchiválás jelenleg nem zajlik, és az államigazgatási, közigazgatási webarchiválásnak is hiányoznak a stratégiai keretei. A 2021-ben hatályba lépett jogi szabályozás az OSZK feladatává teszi a tevékenységet azzal együtt, hogy az állami és önkormányzati szervek, valamint általuk fenntartott intézmények archivált webhelyeit felhasználói engedélykérés nélkül lehet szolgáltatni. Az Országos Széchényi Könyvtárban a kísérleti projekt kereteinek meghatározásakor munkatársaimmal együtt arra a meggyőződésre jutottunk, hogy a központi államigazgatás speciális jellegéből adódóan mindenképp szerencsés lenne, ha ehhez a szférához szorosan kötődő, az egyéb típusú államigazgatási anyagok archiválásában szakértelemmel bíró intézmény vállalná velünk együtt ezt a feladatot. Partnerünk lehetne így a Magyar Nemzeti Levéltár, illetve az Országgyűlési Könyvtár is. Újabb példáját láthatjuk tehát az együttműködés fontosságának, hogy milyen fontos a teljes közgyűjteményi

összképet a kutatás során feltárni. A nyilvános szolgáltatás viszont segíthet esetleg majd annak áthidalásában, hogy hozzá lehessen illeszteni egy adott közigazgatási intézmény archivált tartalmainak nagyját a más intézményben tárolt webes tartalmakhoz.

Sajnos a magyar digitális múzeumi terület különféle tevékenységeinek sem képezi a webarchiválás jelenleg a részét. A múzeumok jelenleg nem tartják feladatuknak, hogy a webes kulturális örökség gyűjtőkörükbe eső részét feltárják, dokumentálják, archiválják, ezzel pedig a jelenünk egy fontos szeletevész el a jövő számára. Remélem, hogy a jövőben egyre több múzeumi partnert is be lehet majd vonni a muzeológiai alapú webarchiválási tevékenységbe.

A világ számos országában, így Magyarországon is, a film, illetve a lineáris média archiválásának keretei már régóta adottak. A médiakörkép azonban az utóbbi évtizedekben jelentősen kibővült. Ma már a hagyományos lineáris tartalomszolgáltatással egyenrangúvá lett a különféle közösségi videomegosztó platformok által szolgáltatott tartalom. Ez a folyamat az utóbbi években publikációs szinten is szépen nyomon követhető (Acker & Kreisberg, 2019; Elson, 2020; Espley és mtsai., 2014; Thomson & Kilbride, 2015). Fontos jellemzője az online tartalomszolgáltatásnak a belépési küszöb hiánya. A közösségi platformok (pl. YouTube, Facebook) megjelenésével már bárki könnyen megjelenhet az audiovizuális virtuális térben is, s e tartalmak szerves elemévé váltak a populáris kultúrának. Mint a brit kormányzati webarchívum esetében (National Archives, 1996) láttuk, immár az államigazgatás (de általában a társadalmi, kulturális intézmények) társadalmi kommunikációjának is egyre megkerülhetetlenebb elemét alkotják. De maguknak a lineáris médiaszolgáltatóknak is markáns online jelenléte van a közösségi platformokon is. Külön kategóriát jelentenek az audio, illetve video podcastok, melyek sajátos átmenetet képeznek a lineáris, illetve az online műsorszórás világa között, hiszen a virtuális térbe sugárzott, letölthető, offline is fogyasztható médiatartalmakról van szó. Világossá kell, hogy váljon mindenki számára, hogy a világháló ezen szelete is fontos külön szeletét alkotja kulturális örökségünknek. Véleményem szerint az online audiovizuális tartalmak archiválása természetes kiegészítő tevékenységeként kell, hogy megjelenjen a hagyományos médiaarchiválási tevékenységeknek. Másfelől viszont a kutatás során világossá vált számomra,

hogy a feladat, illetve a dokumentumtípus speciális jellege miatt ezt a feladatot a könyvtárak, levéltárak vagy múzeumok általános webarchiválási gyűjtőköri feladatként nem is tudják felvállalni. Van néhány olyan ország, például Dánia (Schostag & Fonss-Jorgensen, 2012) ahol a gyűjtőköri és intézményi keretek megfelelően rögzítésre kerültek, sőt Sabine Schostag felhívta a figyelmet arra, hogy külön útmutató készül például YouTube anyagok archiválására (Schostag, 2019a). Ott szerencsés felállás, hogy a médiaarchívum is a nemzeti könyvtár része, így egységes keretek között tudják kezelni a különféle digitális megőrzési tevékenységeket. Nálunk a 2021. január elsejétől hatályos jogi szabályozás az audiovizuális területet épp annak különleges volta miatt kiemelte az általános webarchiválás szabályozási keretéből, s külön jövőbeni szabályozással a Nemzeti Audiovizuális Archívumhoz (NAVA) rendelné ezt a feladatot. A megfelelő szakmai, infrastrukturális és finanszírozási feltételek megteremtése azonban elengedhetetlen ahhoz, hogy a NAVA tevékenysége erre az intézmény számára természetesnek tűnő irányba is kibővüljön.

6.2 A webarchiválás magyarországi gyakorlata kezdetei feltárásának főbb elemei

A szervezett formában elinduló webarchiválás itthoni előzményeinek áttekintése után áttekintést adok a dolgozatban arról az eredeti koncepcióról, melyet a kísérleti projekt munkálatainak kezdeteikor munkatársaimmal közösen felállítottunk. Ezt követően rátérek arra, hogy a gyakorlatban ehhez képest milyen munkafolyamat rendszert alakítottunk ki a webarchiválás megalapozása kapcsán. Sarkalatos pontja tevékenységünknek a gyűjtőköri szabályozás, illetve a gyűjteményfejlesztési irányelvek kijelölése is. Ezt követően kerül bemutatásra az a hardveres, illetve szoftveres háttér, melyre építve alkottuk meg a kísérleti projekt informatikai kereteit. Nagyon fontos összetevője tevékenységünknek a szelektíven aratott és válogatott állomány megfelelő metaadatsémákon keresztüli leírása, feltárhatóvá tétele, ennek a munkának a kezdeteiről is áttekintéssel szolgálók. Lényeges, önálló kutatások tárgyaként szolgáló tényezőként merül fel továbbá a (főként válogatott módon learatott) tartalmak minőségellenőrzésének kérdésköre, a webarchiválást érintő taxonómiai keretek taxonómia elméleti és gyakorlati háttere. A kutatási hipotézisekben rögzített témaköröket alapul véve számos megfontolásra érdemes kérdés merül fel nem csupán gyűjtőköri, hanem

technikai és közgyűjteményi szakmai, szolgáltatásmenedzselési értelemben is arról, hogy mely vonatkozásokban lehet megállapítani a webarchiválási tevékenységek hatókörének határait. Elengedhetetlen kialakítani a jelen műszaki és szakmai jellegű kompetenciáinkhoz kötött hosszútávú megőrzési kereteket is. Sok nyitott kérdés merült fel a magyarországi webarchiválási tevékenység jogi hátterének rendezése kapcsán, mely szintén áttekintést kíván. Végül, de nem utolsósorban pedig arról számolok be, hogy a projekt szakmai megalapozásához, a kutatás-fejlesztési eredmények ismertetéséhez milyen kommunikációs tevékenységeket rendeltünk munkatársaimmal együtt hozzá.

A tézisek kapcsán már korábban kifejtett nagyon fontos része az elemző tevékenységnek a webarchiváláshoz kapcsolódó oktatási irányok, tevékenységek összegzése, valamint a webarchívumok jellemzése a tudományos kutatások tárgyaként.

Kapcsolódó előadások, publikációk

- Drótos, L., & Németh, M. Az OSZK-ban folyó kísérleti webarchiválási projekt első évének tapasztalatai. Tudományos és Műszaki Tájékoztatás, 65(7–8), 389–400, 2018
- Németh M. Az OSZK Webarchívum új honlapjának felépítése és szolgáltatásai, Könyv, Könyvtár, Könyvtáros 29(6), 16-26, 2020
- Drótos, L. & Németh, M. Beszámoló az OSZK webarchiválási pilot projektjéről, valamint a nemzetközi tapasztalatokról. OSZK Könyvtártudományi Szakkönyvtár, Budapest, 2017
- Drótos, L. & Németh, M. Hungarian web archiving pilot project in the National Széchényi Library. Demo paper. 8th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Debrecen, 2017
- Drótos L. & Németh, M. Az OSZK-ban zajló webarchiválási projekt első évének tapasztalatai. Fórum Networkshop 2018 konferencia, Eger, 2018

[T7] Az Országos Széchényi Könyvtár webkönyvtárosaként a szerző meg kívánja alapozni a webarchiválás általános elméleti kutatási hátterét, valamint tevékeny részt vállal az összes tárgyalat webarchiváláshoz kötődő munkafolyamat gyakorlati meghonosításában is.

7.1 A magyar nemzeti könyvtári webarchiválási szolgáltatás fő szolgáltatási koncepciója

Az elméleti háttér főbb elemei a korábbi tézisek kapcsán már kifejtésre kerültek. A webarchiválás eredendő töredék alapúságát mindig szem előtt kell tartanunk. Az egyes közgyűjteményi ágakhoz tartozó intézmények gyűjtőkörének meghatározására megfelelő nemzetközi példák állnak rendelkezésre, ezek feltárása is fontos részét képezte a dolgozathoz kötődő kutatási tevékenységnek. A legnagyobb tétje az előbb vázoltak felmérése mellett mégis csak annak volt, hogy az elméletben felvázoltakat, a nemzetközi gyakorlati elemeket miként lehet a magyar gyakorlatra úgy lefordítani, hogy ne csupán kísérleti szinten foglalkozhassunk az OSZK-ban a webarchiválással, hanem abból állandó könyvtári szolgáltatás is születhessen.

A kísérleti kutatás-fejlesztés céljaként egy leendő magyar internet archívum (webarchívum) koncepciójának, szervezeti kereteinek és informatikai infrastruktúrájának megteremtését, valamint egy tesztarchívum felállítását foglalmaztuk meg munkatársaimmal együtt. Ennek segítségével tapasztalatok gyűjtését kívántuk előtérbe helyezni egy majdani nagy kapacitású, üzemszerűen működő rendszerhez. Végző célként egy olyan rendszer létrehozatalának kívánalma fogalmazódott meg, amely az interneten nyilvánosan megjelenő magyar és magyar vonatkozású kulturális örökség hosszú távú megőrzésének feladata mellett képes kiszolgálni az oktatás, a tudományos kutatás, az állami szervek, az üzleti szféra és az egyes internet használók igényeit is (Drótos, 2017c). A rendszer megteremtésének fő indokának azt jelöltük meg, hogy az addig csupán jelen időben létező magyar internetnek „múltja” is legyen. Olyan lehetőségek megnyitását helyeztük előtérbe a jelen és a jövőbeni időszakok felhasználói számára, amelyek egy webarchívum hiányában nem, vagy csak nehézkesen valósíthatók meg (pl. megszűnt weboldalak megtalálása, régi vagy jelenkori, de feltört ill. véletlenül törölt webhelyek helyreállítása, weboldalak időbeli változásának elemzése és vizualizálása, stabil hivatkozhatóság, idődimenziót is tartalmazó szöveg- és adatbányászati alkalmazások futtatása, internettörténeti kutatások, hiteles másolatok szolgáltatása, igény szerinti webarchiválás megrendelésre történő szolgáltatása) (Drótos, 2017c). A rendszert a külföldi legjobb gyakorlatok tapasztalatainak felhasználásával terveztük kialakítani,

együttműködésben a hazai közgyűjteményi és informatikai szféra egyes szereplőivel. Fő célkitűzésként szabtuk meg, hogy alkalmas legyen nagy tömegű webaratásra, egyes webhelyek vagy weboldalak szelektív lementésére, valamint online tartalmak önkéntes beadására is (Drótos, 2017c). Az ehhez szükséges infrastruktúra megteremtése mellett rögzítésre került a webarchiválás és az archivált anyag metaadatolásának módszertana kialakításának kívánalma; az archívum és az eredeti tartalomgazdák, valamint az archívum és annak felhasználói közötti jogviszonyt szabályozó szerződésminták előállításának is célként fogalmazódott meg; továbbá ajánlások, tanulmányok, előadások és tanfolyamok segítségével kívántuk megismertetni az érintett szereplőkkel a webarchiválással és a webarchívum-használattal kapcsolatos ismereteket (Drótos, 2017f).

A koncepció általános céljai mellé a kutatás fejlesztési-munka konkrét formába öntésére egy 13 tételből álló feladatlista is hozzá lett rendelve, melyeknek végrehajtása 2018 végéig lett előírányozva (Drótos, 2017f). Ezáltal alkalom nyílik a kutatás során annak felmérésére is, hogy az előírányozott célok mennyiben valósultak meg. Az egyes tételek sorrendje egyben időbeli sorrendet is jelöl, az egyes feladatok teljesítésének sorrendjét illetően. Két fő feladatot és tizenegy azokhoz kötődő háttérfeladatot határoztunk meg.

1. Tesztarchívum létrehozása. Az elméleti ismeretek megszerzése mellett gyakorlati tapasztalatok megszerzésének céljából felállítani terveztünk egy kezdetben néhány száz (főként kulturális és tudományos) webhelyből, valamint az Elektronikus Periodika Archívumban csak távoli forrásként nyilvántartott kb. 1700 elektronikus periodika, néhány alkalommal történő mentéséből álló tesztarchívumot. Ez sikeresen meg is valósult. Továbbá meg kívántuk kísérni a .hu domain alatti webszerverek kellően nagy halmazának learatását egyszeri vagy kétszeri alkalommal (mivel egy ekkora aratásnak a rendelkezésre álló infrastruktúrától függően akár több hónapig történő futása is előre jelezhető volt az azt megelőző külföldi tapasztalatok szerint). Ezek a teszt jellegű, terheléses próbák is sikerrel zárultak, segítettek felmérni a domain szintű aratás infrastruktúra szükségleteit. Komolyabb zárt vagy nyilvános hatókörű szolgáltatást nem terveztünk kialakítani erre a gyűjteményre, egyrészt a szerzői jogi háttér hiánya miatt, másrészt pedig, miután a szoftveres infrastruktúra folyamatos beállítását, finomhangolását is el kívántuk végezni, ezért az első

néhány aratásból nem is vártunk szolgáltatásban felmutatható eredményeket. Egy kis demó gyűjteményt azonban a szerzői jogi engedélyek megszerzését követően a jó minőségű mentésekből közzé kívántunk tenni pl. leartott könyvtári webhelyekből, bemutatva a webarchívum lehetséges szolgáltatási alapfunkcióit. - Erre a kísérleti projekt honlapjába beágyazva sor is került, a projekt végeztéig pedig a gyűjtemény folyamatosan bővült újabb kategóriákkal és tételekkel.

2. A tesztarchívum létrehozása mellett el kívántuk kezdeni a szelektíven (is) archiválendő webhelyek vagy webhelyrészek kellően reprezentatív méretű halmazát is feltérképezni. Egy mintegy tízezres nagyságrendű URL-lista kialakítását határoztuk meg, kiegészítve néhány további metaadattal (pl. a webhely neve, fenntartója, műfaja, témaköre, mélysége, változékonysága, veszélyeztetettsége, arathatósága, jogi előírásai, a tartalomgazda elérhetősége). Ezt az adatbázist jelöltük meg a majdani üzemszerűen működő archívum gyarapítási forrásaként, ez alapján terveztük indítani - optimális gyakorisággal és konfigurációs beállításokkal - az egyes webhelyek mentéseit. A címlista bővítésébe és karbantartásába a projekt iránt érdeklődő partnerintézmények munkatársait is be kívántuk kapcsolni, sőt a magyar internetes társadalom szélesebb rétegeinek figyelmére is számítani kívántunk (felhasználókéra és szolgáltatókéra egyaránt), akiknek lehetővé kívántuk tenni, hogy szintén javasolhassanak majd megőrzésre érdemes tartalmakat. - Maga az alapgyűjtemény létrejött, s most már az állandó szolgáltatás részeként bővül dinamikusan újabb és újabb témakörökkel, valamint a tematikus alapú ütemezett aratások anyagával.

3. A külföldi jó gyakorlatok megismerése, a téma aktuális szakirodalmának feldolgozásával együtt. Ennek eredményeit a dolgozatban önálló fejezetben foglaltam össze.

4. Csatlakozás a webarchívumokat tömörítő nemzetközi konzorciumhoz (IIPC). A csatlakozásra 2018 elejétől sor került, az előnyökről, eredményekről is szó esett már.

5. A magyar webarchívum gyűjtőkörének meghatározása. - Egy javaslat készült el ennek kapcsán, melynek véglegesítése és hivatalos formába öntése még nem történt meg.

6. A magyar webtér lehatárolási szempontjainak felmérése. - A felmérés sikeresen lezajlott, alapjául szolgál a domain szintű aratások tervezésének.

7. A metaadatstruktúra kialakítása a szelektív ill. a tömeges mentések esetében is, és a szelektív mentések folyamatos metaadatolási munkafolyamatának kialakítása. - Az XML alapú metaadatséma kialakításra került. Az RDA alapú metaadat séma megteremtése még folyamatban lévő feladat.

8. A tartalomszolgáltatókkal kötendő szerződés tartalmának és az engedélyt kérő e-mail szövegének megfogalmazása. - A munkafolyamat kialakításra került. A szükséges emberi erőforrás hiánya az adminisztráció terén fokozottan jelentkezik, hátráltatva a szabadon szolgáltatható jogvédett állomány dinamikusabb gyarapítását is.

9. Az archiválási tevékenység és az archívumhoz való hozzáférés jogi keretének megteremtése. A jogi keretek 2020. január elsejei hatállyal megteremtődtek.

10. A kísérleti időszak tapasztalatai alapján az üzemszerű működésre alkalmas, jól skálázható informatikai infrastruktúra megtervezése. - Több hullámban került sikeresen megtervezésre, ennek néhány alapvető vonását már kifejtettem egy korábbi tézishez kapcsolódva.

11. A webarchiválással, a webarchívumok működtetésével, használatával kapcsolatos ismeretek oktatása az OSZK, illetve egyéb könyvtárak, közgyűjtemények érdeklődő munkatársai számára. - Belső tanfolyamokat, tájékoztatókat tartottunk, illetve 2018-ban elindult az államilag akkreditált könyvtáros továbbképzési program is a Könyvtári Intézettel együttműködve. Ennek részletei az oktatással kapcsolatos fejezetben kerülnek kifejtésre.

12. A magyar online tartalomszolgáltatók és a sajtó folyamatos tájékoztatása a tervekről és a projekt előrehaladásáról. - A kommunikációs keretek kialakítása megtörtént, amelyet a dolgozatban részletesen kifejttek.

13. A kísérleti időszakot követő állandó szolgáltatás fenntarthatóságának megtervezése. -Az állandó szolgáltatás keretei jogilag is alátámasztva kialakításra kerültek, 2020. január elsején történt indulással.

A kísérleti projekt kereteinek kialakításakor már a kezdetektől építeni kellett a társ digitális gyűjteményekben (Magyar Elektronikus Könyvtár, Elektronikus

Periodika Archívum) felhalmozott tapasztalatokra, a gyarapítási munka során történő együttműködésre.

Az eredeti koncepció mellé számba vettem a személyi feltételeket (humán erőforrás) is. Az osztályvezető mellett megneveztük a kísérleti projekt összehangolását, irányítását gondozó témafelelős pozícióját, ezt a feladatkört a kezdetektől Drótos László látja el, legalább egy webkurátort, aki a gyűjtemény gondozásáért, az ezzel összefüggő jogi-adminisztratív teendők ellátásáért felel, Visky Ákos László dolgozik most ebben a pozícióban. További munkatársakat is szeretnénk bevonni a kurátori, jogkezelési tevékenységekbe. A webkönyvtáros feltárja a nemzetközi jó gyakorlatokat, a téma szakirodalmának összegyűjtését gondozza, részt vesz a nemzetközi kapcsolatok alakításában, a projekt szakmai tervezésében, az oktatási tevékenységek megalapozásában. Ezt a pozíciót töltöm be a kezdetektől fogva, jelen pillanatban is. Ez a webkönyvtárosi tevékenység adott gyakorlati kereteket számomra a jelen disszertációt megalapozó kutatás-fejlesztési tevékenységnek. Informatikai oldalról a hardver és szoftverstruktúra kiépítéséért és üzemeltetéséért felelős rendszergazda, illetve egy alkalmazásfejlesztési és optimalizálási feladatokat ellátó programozó felvételét vettük tervbe. Ebből annyi valósult meg, hogy egy belső és egy külső foglalkoztatásban álló kolléga (Kovács Péter, illetve Vitéz Gábor) segíti negyed- illetve félállásban a munkánkat, napi szinten leginkább utóbbi segítségére tudunk támaszkodni.

Kapcsolódó előadások, publikációk

- Drótos, L., & Németh, M. Metadata Management and Future Plans to Generate Linked Open Data in the Hungarian Web Archiving Pilot Project. ITlib, 2019 (2), 38-44.
- Németh M. Az OSZK Webarchívum új honlapjának felépítése és szolgáltatásai, Könyv, Könyvtár, Könyvtáros 29(6), 16-26, 2020
- Drótos L. & Németh, M. How to catalogue a web archive? Information Interactions (INFINT) 2018 Conference. Department of Library and Information Science, Faculty of Arts, Comenius University & University Library, Bratislava 2018
- Németh M. A short introduction to the Hungarian Web Archiving Pilot Project. IIPC General Assembly, Zagreb, 2019
- Németh, M. The first two years of the Hungarian Web Archiving Pilot Project. Colloquium of Library and Information Experts of the V4+ Countries, Bratislava, 2019

[T8] A webarchiválás általános áttekintése mellett saját kutatási témák is megfogalmazásra kerülnek (például a szemantikus web és a webarchiválás összekapcsolódása révén). Folyamatosan fejlődő dinamikus tudományterületről van szó, ahol csupán pillanatnyi összkép felvázolására nyílik mód, az egyes résztémák alaposabb tanulmányozása folyamatos tudományos feladatot jelent majd a jövőben is.

8.1 Webarchiválás, mikroadatok és a szemantikus web

A mikroadatok webarchiválási célú felhasználásának felvázolásáról a röviden már szó esett a korábbiakban. A schema.org jelölőkkel rendelkező HTML oldalak sokkal értelmezhetőbbek a keresőmotorok számára, mint az azokat nélkülöző társaik, s így tartalmilag sokkal relevánsabb, pontosabb információkhoz lehet hozzájutni általuk (Németh, 2019c). Ezek a jelölők voltaképpen szabványos mikroadatok, melyek az adott oldal forráskódjában jelennek meg (Adamich, 2015; „Getting started with schema.org using Microdata”, 2016). A mikroadatok ugyanakkor nem csupán az indexelés pontosságának növelése érdekében hasznosak. A kapcsolt adatok univerzumában a szabványos szemantikus formában leírt (címkézett) adatkészletek összekapcsolhatók egymással, olyan hatalmas tartalmi bázist kínálva az információk visszakeresése számára, mellyel eddig az emberiség történetében nem találkozhattunk (Németh, 2019c). A mikroadatok által az adott weboldalba, webhelybe beépíthető egy olyan információkészlet is, mely azok hosszútávú digitális megőrzéséhez járul hozzá (Enis, 2015; „Getting started with schema.org using Microdata”, 2016). Az OCLC közrebocsátott egy útmutatót, mely a webarchiváláshoz szükséges leíró jellegű metaadatok alapkészletét teszi közzé (Jackie & Bowers, 2018). Ebben hozzáillesztették a megfelelő schema.org megfeleltetéseket a főleg Dublin Core metaadat rendszerből kiinduló alapelemekhez. Ez az ajánlás azonban csak a lehető legszükségesebb leíró elemeket tartalmazza. További kihívás, hogy a schema.org vagy akármely egyéb szemantikus szótár nem tartalmaz olyan speciális adatelemeket, melyek a webarchiválási tevékenységekhez kapcsolódnak.

A szemantikus web, illetve a webarchiválás összeházasításának egyfajta új módszerét vázolom fel a dolgozatban a szemantikus alapú mikroadatok felhasználási lehetőségei kapcsán. Néhány típusú mikroadat segítséget nyújthat számunkra abban, hogy az arató robotok számára pontosan és teljesen feltérképezhető legyen egy adott webhely szerkezete. A webhely térképe elkészíthető XML formátumban. Ezt a leírást egy robots.txt nevű fájlban tárolható el, melybe speciális utasítások illeszthetők melyek a webarchiválást végző robotok tevékenységét befolyásolják. A robots.txt fájlban tárolt információk megjelenése ki is bővíthető, hogy ne csupán egyetlen fájl jelenjen meg az adott webhely gyökérkönyvtárában, hanem egyes információk a webhely egyes oldalainak fejlécében legyenek rögzítve, pontosan leírva az adott oldalon található adatokhoz történő hozzáférés módjait. Ezek a robots.txt fájlban, illetve az egyes lapok fejléceiben rögzített járulékos információk jelenleg nem szabványos kiegészítések formájában érhetőek el, ilyen például a Robots Exclusion Standard („Robots exclusion standard—Wikipedia”, 2020). Ebben vannak leírva például a webhely egyes oldalainak, illetve elemeinek indexelését kizáró noindex metaadat címke, vagy noindex http választ nyújtó fejléc elem (html response header). De egyéb utasításokat is adhatunk a robotok viselkedésének szabályozására, mint például az allow a sitemap host, vagy a universal match. A lényeges problémát itt az jelenti, hogy ez a szabvány, illetve annak elemei nem részei a hivatalos html szabvány leírásának. A webarchiválást végző szakmai közösségnek arra kellene törekednie, hogy bekerüljenek abba, s elérhetőek legyenek minden robot számára az összes platformon, a lehető legtöbb webhelyen. Ennek segítségével sokkal pontosabb honlaptérképek készülhetnének el a mostaniaknál. Az öröknaptár alkalmazások, a robotok számára nem bejárható adatbázisokra mutató linkek például mind külön címkézhetőek lennének, hogy kizárásra kerüljenek az aratásból, megelőzve, hogy a robot olyan csapdába kerüljön, melyből nem tud kijutni (Németh, 2019c). A noindex http fejléc elem pedig adott esetben egy webhely egész oldalát is kizárhatja az aratásból, amennyiben annak tartalma nem aratható le a robotok számára. A robotok működését szabályozó szabvány további kiterjesztése lehetne olyan szabályok kidolgozása, melyek az adott webhely különféle platformokhoz illeszkedő felületeit határoznák meg (mobilbarát felület, akadálymentes felület, robotbarát felület stb.). Így meg lehet adni a robotok

számára, hogy csupán egy bizonyos felületet járjanak be, vagy külön-külön akár az összeset is, ahogy az archiválási szabályok megalkotásakor rendelkezünk az adott webhely szerkezeti, formai és tartalmi jellemzőinek ismeretében (Németh, 2019c).

Különleges szabályokat állíthatnánk be a robotok számára, amikor a webhely tulajdonosa azt szeretné, hogy csak egy bizonyos felület legyen aratható. Ugyanígy válogathatna a webhely tulajdonosa a különféle nyelvi verziók arathatóságának beállítása kapcsán is. Külön kihívást jelent, amikor egy webhely kezdőoldala csupán valamely egyszerű lekérdező elemet tartalmaz, melyről tovább haladva érhetőek el különféle műveletekkel az egyéb tartalmak. A robot azonban automatikusan nem tudja értelmezni az itt felkínált műveleteket, ezért meg kell adni számára azt a kezdőpontot, esetleges alternatív felületet, ahonnan kiindulva el tudja kezdeni egy adott honlap aratását. Ezt is be lehetne állítani megfelelő mikroadat címke segítségével, így a robotok a kezdőlapról automatikusan átirányításra kerülnének arra az oldalra, ahonnan az aratás elkezdődhet. Ennek révén, amennyiben nem is archiválható a maga teljességében egy webhely teljes keresési funkcionalitása, illetve információ-visszakeresési szolgáltatásai azok eredeti formájában, a felület mögött lévő adatok arathatók és megőrizhetőek lennének (Németh, 2019c).

A fentebb felsorolt információk mikroadatok segítségével történő szabványos megadása a webaratás hatékonyságát még a jelenlegi szoftveres háttérre támaszkodva is jelentősen megnövelné, anélkül, hogy fejlesztői oldalon jelentős befektetéseket kellene eszközölni az archiváló szoftverek továbbfejlesztésének irányában. A webhely tulajdonosoknak, fejlesztőknek pedig a szabványos ajánlások érvényesítése szintén nem jelentene jelentős többlet ráfordítást sem időben, sem pénzben (Németh, 2019c).

Jelentős kihívásként jelentkezik a webarchiválás terén, hogy amennyiben az archiválni kívánt anyag begyűjtése sikeres volt, az archivált állomány a megjelenítő szoftveres háttér korlátai miatt nem jeleníthető meg helyesen. Különféle mikroadatokban elhelyezett utasítások segíthetnék a leartott tartalom szerkezetének és grafikai formátumának helyes megjelenítését. Például számos esetben speciális JavaScript vagy pdf bővítményeket használ egy adott webhely az azokhoz hozzárendelt beágyazott tartalmak megjelenítéséhez, vagy ugyancsak

gyakran lehulló menük tárják fel egy webhely tartalomkínálatát, szolgáltatásait. Amennyiben mikroadatok segítségével a webhely forráskódjába illetve alternatív módokat állítunk be a fotók, dokumentumok, vagy egy egyszerűsített formájú menüstruktúra megjelenítéséhez, akkor az archivált tartalom tökéletesen megjeleníthető lesz (Németh, 2019c).

Speciális mikroadattípusok nyújthatnak segítséget abban, hogy egy adott webhely forráskódjába illetve feltárják a webhelyen használt különféle tartalom formátumok típusait, verzióit és a kapcsolódó szoftverek jellemzőit. Különösen problémás elemek lehetnek a webarchiválás kapcsán a különféle java appletek, beágyazott Flash alkalmazások és tartalmak. Ezeket igen fontos lenne külön mikroadat címkékkel ellátni, meghatározva azok pontos verziószámát és funkciójait. Ezek a pontos leírások különösen fontosá válhatnak majd a jövőben, amikor az archivált webhelyek tartalmának konverziója, illetve emulációja válik majd szükségessé új megőrzési, illetve szolgáltatási platformokon (Németh, 2019c).

A webarchívumok kutatási célú felhasználásának támogatása nagyon fontos részét képezi a szolgáltatási kínálatnak bármely archiválást végző intézmény esetében. A szemantikus adatok szolgáltatása fontos segítséget jelent a kutatók számára az archívumból történő hatékony információkeresésben, amikor számos nagyméretű adatkészlettel, egyéb tartalmi erőforrásokkal kell dolgozniuk. A különféle adatkészletek közötti kapcsolatok megteremtése a nyílt kapcsolt adatok modelljének segítségével szintén igen fontos kutatástámogató tényező. Így ugyanis kiléphet egy adott kutató az archívum virtuális határai közül, s együtt kezelheti az onnan származó adatkészleteket másfajta szabványos adatforrásokkal is. A legfontosabb eszközök e tekintetben a hivatkozást teremtő relációk, amelyek leíró attribútumként rendelhetők hozzá egy adott hiperhivatkozáshoz, meghatározva annak típusát és/vagy annak kapcsolatát a hivatkozás forrása felől annak célpontja felé (Németh, 2019c).

A szemantikus webben az RDF típusú hivatkozások alapvető jelentőségűek a nyílt kapcsolat adatokra épülő adatkészletekben, meghatározva a kapcsolatot (állítás formájában) az RDF tripletekben, lehetővé téve a gépi úton is értelmezhető állítások automatikus feldolgozását egy gigantikus globális gráf keretei között a szemantikus webben. A nyílt kapcsolat adatokra épülő adatkészletekben az RDF

nyelven rögzített linkek adják meg az értékeket az `rdf:type` property elem segítségével, illetve megadják azok kapcsolatait a már jól ismert szemantikus sémákban, szótárakban található kifejezések, definíciók segítségével. Például személynevek, illetve földrajzi nevek is címkézhetők, illetve szűrhetők különféle névtér sémák s az egyes elemekhez rendelt azonosítók segítségével. A szemantikus wikipédia, a szemantikus könyvtári katalógusok, illetve bármely más szemantikusan publikált adatkészletek névtér elemei is hozzáadhatók kapcsolatként (Németh, 2019c).

Egy másik fontos nézőpont egy adott webhely létrehozatala időpontjának, illetve a legutóbbi változtatások dátumának pontos rögzítése. Nagyon fontos rögzíteni egy adott archivált anyag, mint történeti forrás hitelességét is. Sajnos néhány jelenleg is elterjedten használt visszanező szoftver különböző idősíkokban készült mentések elemeit képes egységes felületen megjeleníteni, csorbítva így az archivált anyagok történetileg hiteles megjelenítését. A webhelyek, weboldalak, vagy annak akár egyes részei létrehozatali, módosítási dátum adatainak rögzítése nagyban segíthet abban, hogy az azonos időpontban mentett tartalmi elemek hiteles módon legyenek megjeleníthetők az archivált állományból (Németh, 2019c).

A mikroadatokat néhány digitális bölcsészeti alapú kutatási tevékenység támogatásában is segítséget nyújthatnak. A digitális filológusok nagyszámú XML alapú elemzőeszközt, illetve eljárást használnak. Az általuk használt eszközöket, módszereket megfelelő tartalmú mikroadatokkal szemantikus sémák, szótárak elemei használata révén is támogatni lehet. Például egy angol nyelvű szövegben megbújó ógörög nyelvű idézet, egy adott személy vagy földrajzi név (mint névtér elem) különféle megnevezései, a többnyelvű névalakok, a történeti írásmódok használata is feltárható ily módon (Németh, 2019c).

Mint a fentiekben láthattuk a webarchiválási tevékenységet és a webarchívumok használatát rendkívül változatos módon lehet megfelelő mikroadatforrások előállításával támogatni. A központi kérdés ott bújik meg, hogy ki állítja össze az adatok rögzítését és használatát előíró ajánlásokat, illeszti hozzá azokat a meglévő szabványkörnyezetekhez, sémákhoz, illetve vállalja fel azok folyamatos gondozását. Világosnak tűnik, hogy elsősorban a jelentős webarchívumokat üzemeltető, digitális dokumentumállománnyal bíró közgyűjtemények, illetve

tartalomipari oldalról a nagy tartalomszolgáltató platformok részvétele tűnik ebben elkerülhetetlennek (ideértve a jelentős blogszolgáltatókat, illetve közösségi média felületeket is). Ezek a szereplők elegendő erőforrással rendelkeznek megfelelő ajánlások, illetve szótárak kidolgozásának segítésére, illetve használatuk elterjesztésére. A kritikus felhasználói tömeg elérése a tartalom előállító oldalon kulcskérdés, hiszen így válnak igazán értelmezhetővé a mikroadatok használatának előnyei. A webarchiválás területén az e területen érdekelt IIPC konzorcium vállalhatja vezető szerepet a megfelelő ajánlások és szótárak kidolgozásában. Partnerségi kapcsolatok építésével pedig hozzájárulhatna, hogy a tartalomipar legfőbb szereplői is felismerjék a mikroadatok implementálásának hasznosságát a saját szolgáltatási környezetek keretei közé (Németh, 2019c).

Nagyon hasznos lenne, ha fellendülnének a mikroadatok potenciális alkalmazásával foglalkozó kutatások a webarchiválás területén, s nemzetközi összefogással a közgyűjteményekre támaszkodva, egyre több tartalomipari szereplőt bevonva elterjedne a kidolgozott ajánlások, szótárak, sémák használata, valahogy hasonlóan, mint ahogy az akadálymentes webes felületek kialakításának fontossága is a közgondolkodás részévé vált. A webes kulturális örökség megőrzése ugyanis közös ügye kellene, hogy legyen a tartalomipari szereplőknek s közgyűjteményeknek egyaránt.

Reményeim szerint a jövőben a többi felvázolt kutatási irányban, különösen az adatbányászat, adatvizualizáció kapcsán is jelentős előrelépés lesz tapasztalható. A dolgozat révén talán sikerül ráirányítani a digitális bölcsészek, történészek, adattudósok együttműködésében rejlő potenciálra, s így az OSZK webarchívuma mint a tudományos kutatások tárgya is teret tud nyerni a későbbiekben.

Kapcsolódó előadások, publikációk

- Németh, M., Simon, A.: Public collections on the Semantic web in a Hungarian context. ITLib. 2016 (2), 66-74
- Németh, M.: Semantic web developments in Hungarian public collections in an international perspective. Libellarium. 9 (2), 169-186, 2016.
- Horváth, Á., Németh, M.: Kulturális közkinccsek feltárása nyílt bibliográfiai adatok révén. Könyvtári Figyelő. 63 (2), 196-210, 2017.
- Németh M.: Using semantic microformats for web archiving – an initial project conception. In Katarina Tomková (Eds.) LTP 2019 : Nové trendy a

východiská pri budovaní LTP archívov: zborník príspevkov zo 4. medzinárodnej konferencie o dlhodobej archivácii Bratislava, 5. 11. 2019. (pp. 31-38). Bratislava: Univerzitná knižnica v Bratislave, 2019.

- Németh, M. Webarchívum mint a tudományos kutatások tárgya. Tudományos és Műszaki Tájékoztatás, 67(12), 757–765. 2020.
- Németh M. Potential use of microdata in web-archiving context. CDA 2019. Conference Bratislava, 2019
http://mekosztaly.oszk.hu/mia/doc/CDA_2019-NM.pptx
- Németh, M. A mikroadatok felhasználásának lehetőségei a webarchiválásban. Networkshop 2020 konferencia, Pécs 2020

A kutatás értékelése

Először a hipotézisekben foglaltakkal kapcsolatos kutatás eredményekre térek ki. Az első hipotézisben foglalt célkitűzések véleményem szerint teljesültek. Többféle megközelítésből próbáltam alátámasztani a webarchiválás mint tudományos téma relevanciáját, különös gondot fordítottam annak alapvető közgyűjteményi kereteinek felvázolására. Több fejezetben is szóba kerültek azok a dilemmák, hogy milyen szakmai, technológiai, illetve elméleti módszertani határai vannak a webarchiválással kapcsolatos tevékenységeknek, milyen értelmezési tartományban határozhatjuk meg a webarchiválás fogalomkörét. Önálló témakörként került tárgyalásra ennek kapcsán az archivált állomány minőségének ellenőrzésével, illetve hosszú távú megőrzésével kapcsolatos kihívások halmaza is. Vizsgálódásaim során különös gondot kívántam fordítani a magyarországi webarchiválással kapcsolatos gyakorlati munka, illetve elméleti kutatási háttér megalapozása kapcsán a vonatkozó nemzetközi szakirodalom széleskörű feltárására az egyes, a dolgozatban tárgyalt résztémák kapcsán. A dolgozat fontos összetevőjét alkotja a webarchiválásra irányuló jelenlegi tudományos kutatási irányok összegzése. A webarchiválás mint átfogó interdiszciplináris szakmai kutatási terület egyre jobban intézményesül a nemzetközi tudományos életben. Ugyanez magyar viszonyok között még nem mondható el. Miután a webarchiválás mint állandó szakmai feladat törvénymódosítás révén az Országos Széchényi Könyvtár alaptevékenységei között kap helyet, remélhetőleg a kialakuló gyűjtemény, mint a tudományos vizsgálatok tárgya is szélesebb körben teret nyer majd a későbbiekben. Ezzel a disszertációval voltaképpen a kezdőlökést szerettem volna megadni e folyamat

kibontakozásához. Egyrészt szükséges volt egy általános áttekintéssel szolgálni a webarchiválással kapcsolatos általános tudományos vonatkozásai pontokról. A legfontosabb célja mindezek mellett e disszertációnak mégis az, hogy áttekintést nyújtson a webarchiválást érintő különböző kutatási, illetve kutatás-fejlesztési tevékenységekről. Amellett, hogy meg kívántam jelölni a személyes érdeklődésem homlokterében álló kutatási területeket, egyben inspirációval is kívántam szolgálni ahhoz, hogy minél többen, minél több nézőpontból válasszák a webarchiválást, a webarchívumot mint gyűjteményt a tudományos kutatásaik tárgyául.

Az elméleti és kutatómódszertani keretek mellett a második hipotézisben foglaltak révén a dolgozat másik nagyon hangsúlyos összetevője a formálódó hazai gyakorlat elemeinek felvázolása. A nemzetközi gyakorlat itthon leginkább hasznosítható vonatkozásainak tárgyalása után sort kellett keríteni a tevékenységet jellemző gyakorlati jellegű munkafolyamatok bemutatására. Sokszinű elemzések révén adtam választ arra a hipotézisben feltett kérdésre, hogy sikerrel alapoztuk meg munkatársaimmal együtt a webarchiválást mint önálló nemzeti könyvtári szolgáltatást. A dolgozatban felvázoltam a kiindulási munkatervünket, s azt, hogy ehhez képest meddig sikerült eljutnunk az egyes célkitűzések, munkafolyamatok, szakmai tevékenységek formálása kapcsán. Előkerült az összes olyan főbb tevékenységi irány, melyekkel foglalkozni kellett az állandó szolgáltatás megteremtése kapcsán.

A hipotézisekben foglalt harmadik lényeges témakörként merült fel még a webkönyvtárosi, webkurátori feladatkörök tisztázása, hogy egyáltalán mennyire határozható meg általában a webarchiválás könyvtári feladatként. Ennek kapcsán egyrészt bemutattam a webarchiválás leglényegesebb könyvtári vonatkozásait tágabb közgyűjteményi értelmezési keretek között, másrészt a dolgozatban tárgyalt gyakorlati munkafolyamatok is megfelelő példákat nyújtanak arra, hogy a hagyományos könyvtári tevékenységek, mint például a gyarapítás, feldolgozás, formai és tartalmi leírás, tárolás, megőrzés és a szolgáltatási dimenzió milyen módon értelmezhető a webarchívum mint állományrész vonatkozásában.

A hipotézisekben foglalt negyedik témakör kapcsán nagyon lényeges elemeként került bemutatásra a szabályozási, jogi háttér kialakításának segítése is. Az infrastrukturális, jogi, és a humán erőforrások szűkösségéből adódó, az állandó könyvtári szolgáltatás megalapozása kapcsán felvetődött kihívások időről időre nem kis feladatot jelentettek. A nyitott és többnyire segítőkész nemzeti könyvtári vezetői és a feltétlenül pozitív fenntartói, jogszabályalkotói hozzáállásnak köszönhetően a hipotézisekben felvetett kihívások jó részét sikerült leküzdeni. A webarchiválási tevékenység fenntartható ellátásához azonban feltétlenül szükséges a humán erőforrás bővítése a kezdeti időszakban jelenleg garantált finanszírozási háttér folyamatos fenntartása, az információtechnológiai feltételek folyamatos biztosítása, az intézményes partnerekkel való együttműködés.

Az elméleti háttér és a webarchiválási gyakorlati tevékenység megalapozásának egyfajta közös metszeteként kerültek hangsúlyosan tárgyalásra a hipotézisekben foglalt ötödik témakörben felvetett oktatási kérdések is. A webarchiválás témaköre sikeresen bekerült az akkreditált könyvtáros továbbképzés keretei közé. Lényeges jövőbe mutató feladat e téren a felsőoktatási partnerekkel való kapcsolatok kiépítése, hogy az elméleti és gyakorlati webarchiválási tevékenységeket a graduális képzés oldaláról is támogatni lehessen.

A hipotézisekben foglaltakra adott válaszok összegzése után a publikációs tevékenységem rövid értékelésére kerítek még sort. Minden fő fejezethez természetesen hozzá lettek rendelve a megfelelő hivatkozások révén az önálló publikációk, konferencia előadások. A teljes bibliográfiai adatok pedig a hitelesített publikációs jegyzékben, illetve a dolgozat bibliográfiájában lelhetők fel.

A kísérleti projekt munkafolyamatai kapcsán kiemeltem foglalkoztam előadások, illetve publikációk szintjén a nemzetközi gyakorlattal, a webarchiválás közgyűjteményi kereteinek bemutatásával, a webarchívumok metaadat ellátásával, a webarchiválás oktatásának, tudományos célú felhasználásának kérdéskörével, a szemantikus technológiák használatával, a kísérleti projekt eredményeinek bemutatásával, az állandó nemzeti könyvtári szolgáltatás háttérül létrehozott új honlap tartalomkínálatával (mely egyben keresztmetszetet

nyújt az összes lényeges gyakorlati webarchiválási és azzal összefüggő kommunikációs tevékenységről), valamint a bárki által személyes webarchiváláshoz használható szoftverek áttekintésével.

Végezetül arról szeretnék szót ejteni, hogy az internet fejlődése számos nyitott kérdést rejt, melyek megválaszolása alapvető hatással lehet arra, hogy a weben megjelenő tartalmakat miként lehetséges a jövőben archiválni majd. Megfigyelhető egyfajta egyre gyorsuló széttöredezetttség, egyrészt a különféle platformok szintjén, másrészt a nagyhatalmi rivalizálás virtuális kivetüléseként. Amennyiben a számítógépes világháló egységessége az eddigieknél még jobban háttérbe szorulna, az az archiválás szemszögéből is új helyzetet teremthet.

A webarchiválásnak mint üzleti tevékenységek a tárgya eddig csak nagyon korlátozottan nyert teret. Amennyiben e terület üzletileg esetleg felértékelődne a jövőben, s nagy üzleti súlyú szereplők is megjelennek a szolgáltatásaikkal, az pedig az eddig a döntően a szintéren feltűnő közgyűjtemények, illetve egyéb nonprofit szereplők tevékenységi körének ártértékelésével járhat majd.

Halvány elképzelésekkel rendelkezünk tehát arról, hogy mit hozhat a jövő. Egy azonban biztos. A 2017-ben elindult webarchiválási gyűjteményépítési tevékenység hamarosan szintet fog lépni, s a webarchívum mint gyűjtemény gyorsan bővülő anyaga remélhetőleg egyre több kutató érdeklődését kelti majd fel és sokszínű kutatási együttműködési lehetőségek kialakításának esélyét rejtheti magában.

Bibliográfia

- 1358/2020. (VII. 1.) Korm. Határozat (2020). Letöltés helye: <https://net.jogtar.hu/jogszabaly?docid=A20H1358.KOR&txrefereer=00000001.txt>
2020. Évi XXXII. törvény, Pub. L. No. 32. (2020). Letöltés helye: http://njt.hu/cgi_bin/njt_doc.cgi?docid=219746.383606
- A Kormány 195/2019. (VIII. 1.) Korm. Rendelete (2020). Letöltés helye: <https://magyarkozlony.hu/dokumentumok/b880213d783e1a39ea4b0a3c86290da2ec9c5cac/megtekintes>
- About IIPC. (2019). Letöltés 2020. augusztus 17., helye: <http://netpreserve.org/about-us/>
- Acker, A., & Kreisberg, A. (2020). Social media data archives in an API-driven world. *Archival Science*. <https://doi.org/10.1007/s10502-019-09325-9>
- Adamich, T. (2013). BIBFRAME Take 3: SchemaBibEx, OCLC WorldCat, and Search Optimization. *Technicalities*, 33(5), 9–12. Letöltés helye: <http://search.ebscohost.com/login.aspx?direct=true&db=lxh&AN=97868634&site=eh-ost-live>
- Adamich, T. (2015). Schema BibEx Extensions: Another 21st Century Paradigm Shift. *Technicalities*, 35(4), 15–17. Letöltés helye: <http://search.proquest.com/docview/1726721148?accountid=26452>
- Adobe Flash. (2020). In *Wikipedia*. Letöltés helye: https://en.wikipedia.org/w/index.php?title=Adobe_Flash&oldid=972195466
- Alam, S., Weigle, M. C., Nelson, M. L., & Klein, M.; Van de Sompel, H. (2019). Supporting Web Archiving via Web Packaging, 3. arXiv. 1906.07104. Letöltés helye: <https://pure.know.nl/portal/files/12677017/1906.07104v1.pdf>
- Albanese, A. (2020). Internet Archive Answers Publishers' Copyright Lawsuit. *Publishers Weekly (Online)*. Letöltés helye: <https://www.publishersweekly.com/pw/by-topic/digital/copyright/article/83971-internet-archive-answers-publishers-copyright-lawsuit.html>
- AlSum, A. (2015). Reconstruction of the US First Website. In *Proceedings of the 15th ACM/IEEE-CE on Joint Conference on Digital Libraries—JCDL '15* (pp. 285–286). New York, New York, USA: ACM Press. <https://doi.org/10.1145/2756406.2756954>
- Anderson, K. E. (2020). Getting acquainted with social networks and apps: Capturing and archiving social media content. *Library Hi Tech News*, 37(2), 18–22. <https://doi.org/10.1108/LHTN-03-2019-0011>
- Androvič, A., Bizík, A., Hausleitner, P., Katrincová, B., Lacková, I., & Matúšková, J. (2017). Digitálne pramene—Národný projekt zberu a archivácie v roku 1. *Knihovna PLUS*, (1), 1–14.
- Androvič, A., & Prókai, M. (2007). Web-archívum made in Slovakia: Kísérleti projekt az elektronikus információforrások gyűjtésére és archiválására A web mint kulturális örökség. *Tudományos és műszaki tájékoztatás*, 54(10). Letöltés helye: <https://tmt.omikk.bme.hu/tmt/article/view/6656/7660>
- Archive-friendly website – MIA WIKI. (2020). Letöltés 2020. augusztus 17., helye: https://webarchivum.oszk.hu/mediawiki/index.php?title=Archive-friendly_website
- Archive-It—Web Archiving Services for Libraries and Archives. (2020). Letöltés 2020. augusztus 17., helye: <https://archive-it.org/>
- Asking questions with web archives – introductory notebooks for historians—IIPC. (2020). Letöltés 2020. augusztus 12., helye: <http://netpreserve.org/projects/jupyter-notebooks-for-historians/>
- Az Országos Széchényi Könyvtár gyűjtőköri szabályzata, Pub. L. No. F18/2015 (2015). Letöltés ideje 2020. augusztus 12., helye: http://www.oszk.hu/sites/default/files/f_18_2015_gyujtokori_szabalyzat.pdf
- Bailey, S., Thomson, D., Szalóki, G. (2006). Az első nyilvános webarchívum az Egyesült Királyságban. *Tudományos és műszaki tájékoztatás*, 53(10). Letöltés helye: <https://tmt.omikk.bme.hu/tmt/article/view/5761/6771>

- Belovari, S. (2017). Historians and Web Archives. *Archivaria*, (83), 59–79. Letöltés helye: <https://archivaria.ca/index.php/archivaria/article/view/13600/14985>
- Ben-David, A. (2016). What does the Web remember of its deleted past? An archival reconstruction of the former Yugoslav top-level domain. *New Media & Society*, 18(7), 1103-1119. doi:10.1177/1461444816643790. Letöltés helye: <https://doi.org/10.1177/1461444816643790>
- Ben-David, A. (2020). Counter-archiving Facebook. *European Journal of Communication*, 35(3), 249–264. doi: 10.1177/0267323120922069. Letöltés helye: <https://doi.org/10.1177/0267323120922069>
- Ben-David, A., Amram, A., & Bekkerman, R. (2018). The colors of the national Web: Visual data analysis of the historical Yugoslav Web domain. *International Journal on Digital Libraries*, 19(1), 95–106. <https://doi.org/10.1007/s00799-016-0202-6>
- Beresford, P. (2007). Web Curator Tool. *Ariadne*, (50). Letöltés 2020. augusztus 14., helye: <http://www.ariadne.ac.uk/issue/50/beresford/>
- Berlin, J. (2017). *Wail/Electron*. JavaScript. Letöltés helye: <https://github.com/N0taN3rd/wail> (Original work published 2016)
- Bib.schema.org-1.0—Schema Bib Extend Community Group. (2020). Letöltés 2020. augusztus 14., helye: <https://www.w3.org/community/schemabibex/wiki/Bib.schema.org-1.0>
- Bingham, N. (2014). Quality Assurance Paradigms in Web Archiving Pre and Post Legal Deposit. *Alexandria: The Journal of National and International Library and Information Issues*, 25(1–2), 51–68. doi: 10.7227/ALX.0020. Letöltés helye: <https://doi.org/10.7227/ALX.0020>
- Bingham, N. (2019, július 6). British Library Access Policies: Challenges and Approaches [Presentation]. Letöltés 2020. augusztus 20., helye: <https://digital.library.unt.edu/ark:/67531/metadc1609016/>
- Bratislava University Library. (2016). Webdepozit. Letöltés 2020. január 8., helye: <http://www.webdepozit.sk>
- British Library. (2019). UK Web Archive. Letöltés 2020. január 8., helye: <http://webarchive.org.uk>
- British Library. (2020a). Save a UK website. Letöltés 2020. augusztus 17., helye: <https://www.webarchive.org.uk/en/ukwa/info/nominate>
- British Library. (2020b). UK Web Archive—Technical information. Letöltés 2020. augusztus 17., helye: <https://www.webarchive.org.uk/en/ukwa/info/technical>
- Brozzler – MIA WIKI. (2020). Letöltés 2020. augusztus 18., helye: <https://webarchivum.oszk.hu/mediawiki/index.php/Brozzler>
- Brunelle, J. F., Kelly, M., Salaheldeen, H., Weigle, M. C., & Nelson, M. L. (2015). Not all mementos are created equal: Measuring the impact of missing resources. *International Journal on Digital Libraries*, 16(3–4), 283–301. doi: 10.1007/s00799-015-0150-6. Letöltés helye: <http://dx.doi.org/10.1007/s00799-015-0150-6>
- Brunelle, J. F., Kelly, M., Weigle, M. C., & Nelson, M. L. (2016). The impact of JavaScript on archivability. *International Journal on Digital Libraries*, 17(2), 95–117. doi: 10.1007/s00799-015-0140-8. Letöltés helye: <https://doi.org/10.1007/s00799-015-0140-8>
- Brunelle, J. F., Weigle, M. C., & Nelson, M. L. (2017). Archival Crawlers and JavaScript: Discover More Stuff but Crawl More Slowly. In *Proceedings of the 17th ACM/IEEE Joint Conference on Digital Libraries* (pp. 1–10). Piscataway, NJ, USA: IEEE Press. Letöltés helye: <http://dl.acm.org/citation.cfm?id=3200334.3200336>
- Brügger, N. (2009). Website history and the website as an object of study. *New Media & Society*, 11(1–2), 115–132. doi: 10.1177/1461444808099574. Letöltés helye: <https://doi.org/10.1177/1461444808099574>
- Brügger, N. (2013). Web historiography and Internet Studies: Challenges and perspectives. *New Media & Society*, 15(5), 752–764. doi: 10.1177/1461444812462852. Letöltés helye: <https://doi.org/10.1177/1461444812462852>

- Brügger, N. (2016a). Digital Humanities in the 21st Century: Digital Material as a Driving Force. *Digital Humanities Quarterly*, 10(3), 17. Letöltés helye: <http://www.digitalhumanities.org/dhq/vol/10/3/000256/000256.html>
- Brügger, N. (2016b). Introduction: The Web's first 25 years. *New Media & Society*, 18(7), 1059–1065. doi: 10.1177/1461444816643787. Letöltés helye: <https://doi.org/10.1177/1461444816643787>
- Brügger, N. (2018). *HOW TO DIG INTO THE HISTORY OF A NATION'S WEB? THE DEVELOPMENT OF THE DANISH WEB 2006-2015*. Előadás IIPC RWG webinar series, online webinar. Letöltés ideje 2019. december 15., helye: <https://netpreserve.org/events/iipc-rss-webinar-the-development-of-the-danish-web-2006-2015-2/>
- Brügger, N., Goggin, G., Milligan, I., & Schafer, V. (2017). Introduction: Internet histories. *Internet Histories*, 1(1–2), 1–7. doi: 10.1080/24701475.2017.1317128. Letöltés helye: <https://doi.org/10.1080/24701475.2017.1317128>
- Brügger, N., & Laursen, D. (2019, június 6). A National Web Trend Index [Presentation]. Letöltés 2020. augusztus 20., helye: <https://digital.library.unt.edu/ark:/67531/metadc1608974/>
- Brügger, N., Laursen, D., & Schafer, V. (2019, június 6). Opportunities and Challenges in Collecting and Studying National Webs [Presentation]. Letöltés 2020. augusztus 20., helye: <https://digital.library.unt.edu/ark:/67531/metadc1609008/>
- Brügger, N., & Schroeder, R. (Szerk.). (2017). *The Web as History: Using Web Archives to Understand the Past and the Present* (1. kiad.). United States, North America: UCL Press.
- Cartledge, C., & Nelson, M. (2015). When should I make preservation copies of myself? *International Journal on Digital Libraries*, 16(3/4), 183–205. Letöltés helye: <http://10.0.3.239/s00799-015-0155-1>
- Christiane Lupovici. (2006). Web Crawling: The Bibliothèque nationale de France's Experience. *International Cataloging & Bibliographic Control*, 35(4), 82–84.
- Clifton, G. (2005, március 2). Comments on IIPC web archiving metadata set (v 0.1) [Text]. Letöltés 2020. augusztus 20., helye: <https://digital.library.unt.edu/ark:/67531/metadc1457772/>
- Conifer. (2020). Letöltés 2020. augusztus 18., helye: <https://conifer.rhizome.org>
- Cowls, J. (2015, április 27). Research Using Big UK Domain Data [Presentation]. Letöltés 2020. augusztus 20., helye: <https://digital.library.unt.edu/ark:/67531/metadc1476399/>
- Crawler-friendly website – MIA WIKI. (2020). Letöltés 2020. augusztus 17., helye: https://webarchivum.oszk.hu/mediawiki/index.php?title=Crawler-friendly_website
- Cunningham, A., & Phillips, M. (2005). Accountability and accessibility: Ensuring the evidence of e-governance in Australia. *Aslib Proceedings*, 57(4), 301–317. doi: 10.1108/00012530510612059. Letöltés helye: <https://doi.org/10.1108/00012530510612059>
- Czech National Library. (2015a). Comprehensive Harvest- Czech Webarchive. Letöltés 2020. január 8., helye: <http://www.webarchiv.cz/en/comprehensive-harvests>
- Czech National Library. (2015b). Czech webarchive-topic collections. Letöltés helye: <http://www.webarchiv.cz/en/topic-collections>
- Czech Webarchive. (2011). Webarchiv.cz. Letöltés 2020. január 8., helye: <http://www.webarchiv.cz>
- Dancs, Sz. (2011). Webarchiválási politikák. *Könyv, könyvtár, könyvtáros*, 20(10), 14–20. Letöltés helye: http://www.epa.hu/01300/01367/00248/pdf/EPA01367_3K_2011_10.pdf
- Danmarks Kulturministeriet. (2015). Pligtaflevering lov. Letöltés 2020. január 8., helye: <http://pligtaflevering.dk/loven/index.htm>
- Davis, R. C. (2016). The Future of Web Citation Practices. *Behavioral & Social Sciences Librarian*, 35(3), 128–134. doi: 10.1080/01639269.2016.1241122. Letöltés helye: <http://dx.doi.org/10.1080/01639269.2016.1241122>

- Decman, M. (2011). Problems of Long-Term Preservation of Web Pages TT - Problematika Dolgorodne Hrambe Spletnih Strani. *Knjižnica*, 55(1), 193–208. Letöltés helye: <https://search.proquest.com/docview/1266143497?accountid=27464>
- Deschamps, R. (2019, március 12). Exploring Web Archival Data through Archives Unleashed Cloud Jupyter Notebooks. Letöltés 2020. augusztus 21., helye: <https://news.archivesunleashed.org/exploring-web-archival-data-through-archives-unleashed-cloud-jupyter-notebooks-7605c6ca2b33>
- Di Pretoro, E., & Geeraert, F. (2019). Behind the Scenes of Web Archiving: Metadata of Harvested Websites. In Depoortere, R., Gheldof, T., Styven, D., & Van, J.Eycken, D. (Szerk.), *Press, Trust and Understanding: The value of metadata in a digitally joined-up world* (pp. 63–74). Brussels: Archives et Bibliothèques de Belgique - Archief- en Bibliotheekwezen in België. Letöltés helye: <https://hal.archives-ouvertes.fr/hal-02124714>
- Drótos, L. (2006). Mi a MIA? Javaslat egy magyar Internet Archívum létrehozására. *Tudományos és műszaki tájékoztatás*, 53(6), 267–274.
- Drótos, L. (2017a). Az internet archiválása mint könyvtári feladat. *Tudományos és műszaki tájékoztatás*, 64(7–8), 361–371.
- Drótos, L. (2017b). OpenWayback. MIA WIKI. Letöltés 2020. február 23., helye: <http://mekosztaly.oszk.hu/mediawiki/index.php/Wayback>
- Drótos, L. (2017c). *OSZK webarchívum pilot koncepció -2017-04-17*.
- Drótos, L. (2018a). *Webarchívum „gépház”*. Prezentáció, Budapest. Letöltés ideje 2018. december 20; helye: http://mekosztaly.oszk.hu/mia/doc/Webarchivum_gephaz.pptx
- Drótos L. (2018b). Webes tartalmak digitális megőrzése. *Könyv, könyvtár, könyvtáros*, 27(10), 11–17. Letöltés helye: https://epa.oszk.hu/01300/01367/00307/pdf/EPA01367_3K_2018_10_011-017.pdf
- Drótos L., & Németh M. (2018a). A webarchiválás oktatása = The education of web-archiving. In *NETWORKSHOP 2018 konferenciakiadvány* (pp.31–37). Budapest: HUNGARNET Egyesület. doi: 10.31915/NWS.2018.4 Letöltés helye: <https://doi.org/10.31915/NWS.2018.4>
- Drótos, L., & Németh, M. (2018b). Web museum, web library, web archive The responsibility of public collections to preserve digital culture. In Petrovska, L., Ivāne-Kronberga, B., & Meldere, Z. (Szerk.), *The Power of Reading: Proceedings of the XXVI Bobcatss Symposium, Riga, Latvia, January 2018* (pp. 124–126). Riga: The University of Latvia Press. Letöltés helye: http://blogi.lu.lv/bobcatss2018/files/2018/08/BOBCATSSS_2018_TheProceedings.pdf
- Drótos, L., & Visky, Á. L. (2018a). *Útmutató a szelektíven archivált egyedi webhelyek XML-alapú metaadatolásához*. Letöltés ideje 2020. november 14; helye: https://docs.google.com/document/d/1InGzbRgFk9RnyK_05IKBcud01F-eli2mT03aqCOTIGk/
- Drótos, L., & Visky, Á. L. (2018b). *Útmutató az archivált webhelyekből/weboldalakból álló részgyűjtemények XML-alapú metaadatolásához*. Letöltés ideje 2020. november 14; helye: https://docs.google.com/document/d/1NlgWfzHLS0AzcVEXweuSHBq9-AkvZFbQw_totLoEG7l/
- Drótos, L., Németh M., Visky, Á.L. (2018). Alapozó ismeretek. Letöltés ideje, 2020. január 8; helye: http://mekosztaly.oszk.hu/mia/e-learning/internet-archivalas_1modul.html
- Dynamic HTML. (2020). In *Wikipedia*. Letöltés helye: https://en.wikipedia.org/w/index.php?title=Dynamic_HTML&oldid=970834686
- Dynamic web page. (2020). In *Wikipedia*. Letöltés helye: https://en.wikipedia.org/w/index.php?title=Dynamic_web_page&oldid=967203212
- Egense, T., & Klindt Myrvoll, A. (2018, november 15). SolrWayback: Demo of the SolrWayback search interface, tools and playback engine for WARC's [Presentation]. Letöltés 2020. augusztus 20., helye: <https://digital.library.unt.edu/ark:/67531/metadc1477121/>

- Elson, A. K. (2020). Getting acquainted with social networks and apps: Capturing and archiving social media content. *Library Hi Tech News*, 37(2), 18–22. doi: 10.1108/LHTN-03-2019-0011. Letöltés helye: <https://doi.org/10.1108/LHTN-03-2019-0011>
- Enis, M. (2015). Ending the Invisible Library | Linked Data. *Library Journal*. doi: 10.1016/S0022-3913(12)00047-9. Letöltés helye: [https://doi.org/10.1016/S0022-3913\(12\)00047-9](https://doi.org/10.1016/S0022-3913(12)00047-9)
- Espley, S., Carpentier, F., Pop, R., & Medjkoune, L. (2014). Collect, Preserve, Access: Applying the Governing Principles of the National Archives UK Government Web Archive to Social Media Content. *Alexandria: The Journal of National and International Library and Information Issues*, 25(1–2), 31–50. doi: 10.7227/ALX.0019. Letöltés helye: <https://doi.org/10.7227/ALX.0019>
- Estonian National Library. (2008). About Estonian Web Archiving. Letöltés 2019. január 8., helye: <https://www.nlib.ee/en/content/estonian-web-archive>
- Fansler, C., Gilbertson, K., & Petersen, R. (2014). The Missing Link: Observations on the Evolution of a Web Archive. *Journal for the Society of North Carolina Archivists*, 11(1), 46–59. Letöltés helye: http://www.ncarchivists.org/wp-content/uploads/2017/08/JSNCA_2014_Vol11No1_Fansler-Gilbertson-Petersen.pdf
- Farag, M., Nakate, P., & Fox, E. A. (2016). Big Data Processing of School Shooting Archives. In *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries—JCDL '16* (pp. 271–272). New York, New York, USA: ACM Press. doi: 10.1145/2910896.2925466. Letöltés helye: <https://doi.org/10.1145/2910896.2925466>
- Final report. Asking questions with web archives—Introductory notebooks for historians.* (2020). Letöltés 2020. december 10., helye: <http://netpreserve.org/projects/jupyter-notebooks-for-historians/>
- Finnemann, N. O. (2019). Web Archive. *KNOWLEDGE ORGANIZATION*, 46(1), 47–70. doi: 10.5771/0943-7444-2019-1-47. Letöltés helye: <https://doi.org/10.5771/0943-7444-2019-1-47>
- Flanagan, G. P. (2010). *Digital Preservation and Authentic Legal Information* (SSRN Scholarly Paper No. ID 2463288). Rochester, NY: Social Science Research Network. doi: 10.2139/ssrn.2463288. Letöltés helye: <https://doi.org/10.2139/ssrn.2463288>
- Fritz, S., & Milligan, I. (2018). Archive-It Blog – Analyze your Web Archives at Scale: The Archives Unleashed Cloud. Letöltés 2020. augusztus 21., helye: <https://archive-it.org/blog/post/analyze-your-web-archives-at-scale-the-archives-unleashed-cloud/>
- Getting started with schema.org using Microdata. (2016). Letöltés 2015. január 1., helye: <http://schema.org/docs/gs.html>
- Getting Started—Schema.org. (2020). Letöltés 2020. augusztus 14., helye: <https://schema.org/docs/gs.html>
- Gomes, D., & Costa, M. (2014). The Importance of Web Archives for Humanities. *International Journal of Humanities and Arts Computing*, 8(1), 106–123. doi: 10.3366/ijhac.2014.0122. Letöltés helye: <https://doi.org/10.3366/ijhac.2014.0122>
- Government Web& Data Archive. (2017). Letöltés 2020. január 6., helye: <https://digital2.library.unt.edu/nomination/GWDA/about/>
- Grab Them All. (2017). Letöltés 2020. február 23., helye: http://mekosztaly.oszk.hu/mediawiki/index.php/Grab_Them_All
- Grimshaw, J. (2016). UK Official Publications: Managing the Transition to Electronic Deposit at the British Library. *Legal Information Management*, 16(1), 3–9. doi: 10.1017/S1472669616000037. Letöltés helye: <http://dx.doi.org/10.1017/S1472669616000037>
- Gulyás, L., Jurányi, Z., Soós, S., & Kampis, G. (2014). Can web presence predict academic performance? In *Proceedings of the 23rd International Conference on World Wide Web—WWW '14 Companion* (pp. 1183–1188). New York, New York, USA: ACM Press. doi: 10.1145/2567948.2579037. Letöltés helye: <https://doi.org/10.1145/2567948.2579037>

- Halbert, M., Skinner, K., Wilson, M., & Zardt, F. (2016). Here Today, Gone within a Month: The Fleeting Life of Digital News. In *IFLA WLIC 2016 – Columbus, OH – Connections. Collaboration. Community in Session S21—Satellite Meeting: News Media. In: News, new roles & preservation advocacy: Moving libraries into action, 10-12 August 2016, Lexington, KY, USA*. Lexington, KY, USA: IFLA -- International Federation of Library Associations and Institutions. Letöltés helye: <http://library.ifla.org/id/eprint/2077>
- Hale, S. A., Yasseri, T., Cowls, J., Meyer, E. T., Schroeder, R., & Margets, H. (2014). Mapping the UK Webspace: Fifteen Years of British Universities on the Web. In *Proceedings of the 2014 ACM Conference on Web Science* (pp. 62–70). New York, NY, USA: ACM. doi: doi.org/10.1145/2615569.2615691. Letöltés helye: <https://doi.org/10.1145/2615569.2615691>
- Haskovcova, M. (2019). Webarchiv. Czech National Library. Letöltés ideje 2020. december 14. helye: https://webarchivum.oszk.hu/wp-content/uploads/2020/03/DH_Budapest_2019_Marie_Haskovcova.pdf
- Hegyközi, I. (2014). Hol tart ma a webarchiválás? Szakirodalmi áttekintés. *Könyvtári figyelő*, 60(4), 527–534. Letöltés helye: http://epa.oszk.hu/00100/00143/00095/pdf/EPA00143_konyvtari_figyelo_2014_4_527-534.pdf
- Heritrix—IA Webteam Confluence. (2020). Letöltés 2020. augusztus 18; helye: <https://webarchive.jira.com/wiki/spaces/Heritrix/overview>
- Holub, K., & Rudomino, I. (2015). A decade of web archiving in the National and University Library in Zagreb. In *Preservation and Conservation with Information Technology. IFLA 2015 South Africa*. Cape Town: IFLA -- International Federation of Library Associations and Institutions. Letöltés ideje: 2020 november 15; helye: <http://library.ifla.org/id/eprint/1092>
- Holzmann, H., Goel, V., & Anand, A. (2016). ArchiveSpark. In *Proceedings of the 16th ACM/IEEE-CS on Joint Conference on Digital Libraries—JCDL '16* (pp. 83–92). New York, New York, USA: ACM Press. doi: [10.1145/2910896.2910902](https://doi.org/10.1145/2910896.2910902) Letöltés helye: <https://doi.org/10.1145/2910896.2910902>
- Holzmann, H., Goel, V., & Gustainis, E. N. (2017). *Universal distant reading through metadata proxies with archivespark. 2017 IEEE International Conference on Big Data (Big Data), Big Data (Big Data), 2017 IEEE International Conference on* (p. 459). IEEE. doi: [10.1109/BigData.2017.8257958](https://doi.org/10.1109/BigData.2017.8257958). Letöltés helye: <https://doi.org/10.1109/BigData.2017.8257958>
- Horváth, Á. (2016). *RDFa—Schema.org: Unity of document and semantic web*. Előadás Networkshop 2016, Debrecen. Letöltés ideje: 2020 november 14; helye: <https://conference.niif.hu/event/5/session/10/contribution/27/material/slides/0.ppt>
- How and when to use Brozzler. (2020). Letöltés 2020. augusztus 18; helye: <http://support.archive-it.org/hc/en-us/articles/360000351986>
- HTML 5.2. (2017). Letöltés 2020. augusztus 15; helye: <https://www.w3.org/TR/html52/>
- HTTPS. (2020). In *Wikipedia*. Letöltés helye: <https://en.wikipedia.org/w/index.php?title=HTTPS&oldid=971569450>
- HTTrack Website Copier—Free Software Offline Browser (GNU GPL). (2020). Letöltés 2020. augusztus 18; helye: <https://www.httrack.com/>
- HTTrack—MIA WIKI. (2017). Letöltés 2020. február 23; helye: <http://mekosztaly.oszk.hu/mediawiki/index.php/HTTrack>
- IFLA Library. (2017). Session 186—A memory bank for the digital society: National libraries and web archiving—National Libraries with Information Technology. Letöltés 2020. augusztus 17; helye: <http://library.ifla.org/view/conferences/2017/2017-08-23/798.html>
- IIPC. (2017). OpenWayback. Letöltés 2020. február 23; helye: <https://github.com/iipc/openwayback/wiki>
- IIPC. (2019). IIPC portal. Letöltés 2019. január 28; helye: <https://netpreserve.org>

- IIPC. (2020a). Legal Deposit. Letöltés 2019. január 28; helye: <http://netpreserve.org/web-archiving/legal-deposit/>
- IIPC - Collection development policies. (2018). Letöltés 2020. augusztus 17; helye: <http://netpreserve.org/web-archiving/collection-development-policies/>
- IIPC Training Working Group. (2020). Training materials. Letöltés 2020. augusztus 17; helye: <http://netpreserve.org/web-archiving/training-materials/>
- IIPC -awesome-web-archiving. (2020). IIPC. Letöltés 2020. augusztus 17; helye: <https://github.com/iipc/awesome-web-archiving>
- IIPC Content Development Working Group. (2019). Letöltés 2020. augusztus 17; helye: <http://netpreserve.org/about-us/working-groups/content-development-working-group/>
- IIPC -Országos Széchényi Könyvtár. (2018). Letöltés 2020. augusztus 17; helye: <http://netpreserve.org/about-us/members/orszagos-szechenyi-konyvtar/>
- IIPC, R. W. G. (2020b). Research Working Group—IIPC. Letöltés 2020. augusztus 12; helye: <https://netpreserve.org/about-us/working-groups/research-working-group/>
- IIPC RSS LinkGate Webinar. (2020). Letöltés 2020. augusztus 13; helye: https://docs.google.com/presentation/d/1mYSciOvbU9Hm3jsMSJgioSVr3ZGuVkvYr10HnHXJwXA/edit#slide=id.g8cec5e7f8a_0_67
- IIPC Training Video Case Studies. (2020). Letöltés 2020. augusztus 18; helye: <http://netpreserve.org/web-archiving/iipc-training-video-case-studies/>
- IIPC TWG. (2017). IIPC Training Working Group survey. Letöltés 2018. június 12; helye: <https://www.surveymonkey.com/r/V7MVXXW>
- IIPC TWG. (2019). IIPC Training Working Group. Letöltés 2019. január 28; helye: <http://netpreserve.org/about-us/working-groups/training-working-group/>
- IIPC Web Archiving Metadata Set—UNT Digital Library. (2005). Letöltés 2020. augusztus 20; helye: <https://digital.library.unt.edu/ark:/67531/metadc1457778/>
- IIPC Working groups. (2019). Letöltés 2020. augusztus 17; helye: <http://netpreserve.org/about-us/working-groups/>
- Ilácsa, Sz. (2020). A webarchiválás metaadatolásának problémái. *Könyvtári figyelő*, 30. (66.) (1), 74–82. Letöltés helye: https://epa.oszk.hu/00100/00143/00359/pdf/EPA00143_konyvtari_figyelo_2020_01_074-082.pdf
- International Internet Preservation Consortium (IIPC) repository. (2020). Letöltés 2020. augusztus 2; helye: <https://digital.library.unt.edu/explore/partners/IIPC/>
- Internet Archive. (2019). Internet Archive portal. Letöltés 2019. január 28; helye: <http://www.archive.org>
- Jackie, D., & Bowers, K. (2018). *Descriptive Metadata for Web Archiving: Recommendations of the OCLC Research Library Partnership Web Archiving Metadata Working Group*. Letöltés helye: <https://www.oclc.org/research/publications/2018/oclcresearch-descriptive-metadata/recommendations.html>
- Jackson, A. (2015, április 27). Ten Years of the Uk Web Archive: What Have We Saved? [Presentation]. Letöltés 2020. augusztus 20; helye: <https://digital.library.unt.edu/ark:/67531/metadc1476398/>
- Jackson, A. (2018, november 13). Continuous, Incremental, Scalable, Higher-Quality Web Crawls with Heritrix [Presentation]. Letöltés 2020. augusztus 20; helye: <https://digital.library.unt.edu/ark:/67531/metadc1477144/>
- Jatowt, A., Kawai, Y., Nakamura, S., Kidawara, Y., & Tanaka, K. (2006). Journey to the past. In *Proceedings of the seventeenth conference on Hypertext and hypermedia—HYPERTEXT '06* (p. 135). New York, New York, USA: ACM Press. doi: 10.1145/1149941.1149969. Letöltés helye: <https://doi.org/10.1145/1149941.1149969>
- Jatowt, A., Kawai, Y., Ohshima, H., & Tanaka, K. (2008). What can history tell us? In *Proceedings of the nineteenth ACM conference on Hypertext and hypermedia—HT '08* (p. 5). New York, New York, USA: ACM Press. <https://doi.org/10.1145/1379092.1379098>

- Jatowt, A., Kawai, Y., & Tanaka, K. (2007). Detecting Age of Page Content. In *Proceedings of the 9th Annual ACM International Workshop on Web Information and Data Management* (pp. 137–144). New York, NY, USA: ACM. doi: 10.1145/1316902.1316925. Letöltés helye: <https://doi.org/10.1145/1316902.1316925>
- Java applet. (2020). In *Wikipedia*. Letöltés helye: https://en.wikipedia.org/w/index.php?title=Java_applet&oldid=972322018
- Johnson, P. (2016). Where Did All the Information Go? Well at Least the Important Stuff. *Refer*, 32(2), 8–12. Letöltés helye: <https://search.proquest.com/docview/1803538845?accountid=27464>
- Jupyter notebooks for web archives. (2020). Letöltés 2020. augusztus 12; helye: <https://slides.com/wragge/iipc-jupyter>
- Jurányi, Zs. (2013). Heritrix 3 konfigurációs kézikönyv. Letöltés 2020. február 23; helye: <https://juzraai.github.io/blog/2013/heritrix-3-konfig-receptkonyv/>
- Kampis, G. (2017). *A MIA pilot rövid bemutatása*. Előadás 404 Not Found” workshop, Budapest. Letöltés ideje: 2020 március 14; helye: https://webarchivum.oszk.hu/wp-content/uploads/2020/03/Kampis_Gyorgy_MIA_pilot_GK_ea.pptx
- Kampis, G., & Gulyás, L. (2013). Big is small, and changes slowly in Hungary. In *Coginfo 2013 Conference*.
- Kelly, M. (2020). *Wail—Github*. Python. Letöltés helye: <https://github.com/machawk1/wail> (Original work published 2013)
- Klasinc, J. (2019). Web Archiving Overview: National and University Library - Slovenia [Presentation]. Letöltés 2020. augusztus 20., helye: <https://digital.library.unt.edu/ark:/67531/metadc1609023/>
- Klein, M., Balakireva, L., & de Sompel, H. (2018). Focused Crawl of Web Archives to Build Event Collections. In *Proceedings of the 10th ACM Conference on Web Science* (pp. 333–342). New York, NY, USA: ACM. doi. 10.1145/3201064.3201085. Letöltés helye: <https://doi.org/10.1145/3201064.3201085>
- Klein, M., Shankar, H., Balakireva, L., & Van de Sompel, H. (2019). The Memento Tracer Framework: Balancing Quality and Scalability for Web Archiving. In Doucet, A., Isaac, A., Golub, K., Aalberg, T., & Jatowt, A. (Szerk.), *Digital Libraries for Open Knowledge. TPDL 2019. Lecture Notes in Computer Science, vol 11799*. (pp. 163–176). Cham: Springer International Publishing. doi: 978-3-030-30760-8_15. Letöltés helye: https://doi.org/10.1007/978-3-030-30760-8_15
- Kokas, K., & Drótos, L. (2018). Webarchiválás és a történeti kutatások. *Digitális Bölcsészettudományi Közlemények*, 1(1), 35–55. doi: 10.31400/dh-hun.2018.1.129 Letöltés helye: <https://doi.org/10.31400/dh-hun.2018.1.129>
- Kornhoffer, M. (2010a). *100 nap az élet: A világhálón található információk gyűjtésének és megőrzésének áttekintése hazánkban, valamint Kelet- és Közép-Európában*. Pécs: Pécsi Tudományegyetem. Letöltés helye: <https://redmine.oszk.hu/documents/82>
- Kornhoffer, M. (2010b). *A világhálón található információk gyűjtésének és megőrzésének hazai és nemzetközi áttekintése* (BA szakdolgozat). Pécsi Tudományegyetem. Letöltés helye: <http://mek.oszk.hu/10400/10424>
- Kornhoffer, M. (2011). Internet-archívumok hazánkban és Közép-Európában. *Felderítő Szemle*, 10(3–4), 63–78.
- Köuts, J. (2012). Estonian Web Archive: Preserving the Estonian Mind [Presentation]. Letöltés 2020. augusztus 20; helye: <https://digital.library.unt.edu/ark:/67531/metadc1638385/>
- Kvasnica, J. (2016). Vědecké využití dat z webových archivů. *Knihovnická revue[online]*, 27(2), 23–34. Letöltés helye: <https://knihovnarevue.nkp.cz/archiv/2016-2/recenzovane-prispevky/vedecke-vyuziti-dat-z-webovych-archivu>
- Kvasnica, J., & Kreibich, R. (2013). Formátová analýza sklizených dat v rámci projektu webarchiv NK ČR. (Czech). *File Format Recognition of Data Harvested by Web Archiving Project of National Library of the Czech Republic. (English)*, (2), 1.
- Leach-Murray, S. (2018). Archive-It. *Technical Services Quarterly*, 35(2), 214. Letöltés helye:

- <http://search.ebscohost.com/login.aspx?authtype=ip,cookie,cpid&custid=s6213251&roupid=main&profile=eds>
- Leadership. (2019). Letöltés 2020. augusztus 17; helye: <http://netpreserve.org/about-us/leadership/>
- Lee, J. (2017). Where should the culture of our lives and memory be preserved? - Rethinking the role of the library. In *IFLA WLIC 2017 – Wrocław, Poland – Libraries. Solidarity. Society. In Session 189—Asia and Oceania*. Wrocław: IFLA -- International Federation of Library Associations and Institutions. Letöltés helye: <http://library.ifla.org/1691/>
- Lerner, A., Kohno, T., & Roesner, F. (2017). Rewriting History. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security—CCS '17* (pp. 1741–1755). New York, New York, USA: ACM Press. doi: 10.1145/3133956.3134042. Letöltés helye: <https://doi.org/10.1145/3133956.3134042>
- LinkGate: Core Functionality and Future Use Cases—IIPC. (2020). Letöltés 2020. augusztus 12; helye: <http://netpreserve.org/projects/LinkGate/>
- Littman, J., Chudnov, D., Kerchner, D., Peterson, C., Tan, Y., Trent, R., ... Wrubel, L. (2018). API-based social media collecting as a form of web archiving. *International Journal on Digital Libraries*, 19(1), 21–38. doi: 10.1007/s00799-016-0201-7. Letöltés helye: <http://dx.doi.org/10.1007/s00799-016-0201-7>
- Liu, J. P. (2016). Fair Use, Notice Failure, and the Limits of Copyright as Property. *Boston University Law Review*, 96(3), 833–856. Letöltés helye: <http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=118226656&lang=hu&site=ehost-live>
- Live web leakage – MIA Wiki. (2017). Letöltés 2020. augusztus 18; helye: https://webarchivum.oszk.hu/mediawiki/index.php?title=Live_web_leakage
- Lnenicka, M., Hovad, J., & Komarkova, J. (2015). A Proposal of a Big Web Data Application and Archive for the Distributed Data Processing with Apache Hadoop. In Núñez, M., Nguyen, N. T., Camacho, D., & Trawiński, B. (Szerk.), *Computational Collective Intelligence. Lecture Notes in Computer Science, vol 9330* (pp. 285–294). Cham: Springer International Publishing. doi: 10.1007/978-3-319-24306-1_28. Letöltés helye: https://doi.org/10.1007/978-3-319-24306-1_28
- Locatelli, E. (2017). The role of Internet Wayback Machine in a multi-method research project. In *“Researchers, practitioners and their use of the archived web”, London, School of Advanced Study, University of London*. London.
- Cerbová L. (2009). A cseh kötelezpéldány rendelet és a web. *Könyvtári figyelő*, 55(3), 518–520. Letöltés helye: https://epa.oszk.hu/00100/00143/00072/pdf/2009_3%20szam_referatumok.pdf
- Macnaught, B. (2018). Social Media Collecting at the National Library of New Zealand. In *IFLA WLIC 2018 – Kuala Lumpur, Malaysia – Transform Libraries, Transform Societies Session 93—National Libraries and Social Media—Meeting the Challenges of Acquiring, Preserving and Proving Long-Term Access—National Libraries*. Kuala Lumpur: IFLA. Letöltés helye: <http://library.ifla.org/id/eprint/2274>
- Maeda, N., & Oyama, S. (2017). The Technologies of Web Archiving. *Joho No Kagaku to Gijutsu*, 67(2), 73. Letöltés helye: <https://search.proquest.com/docview/2007445742?accountid=27464>
- Maemura, E., Becker, C., & Milligan, I. (2016). Understanding computational web archives research methods using research objects. In *2016 IEEE International Conference on Big Data (Big Data)* (pp. 3250–3259). IEEE. doi: 10.1109/BigData.2016.7840982. Letöltés helye: <https://doi.org/10.1109/BigData.2016.7840982>
- Masanès, J. (2005). Web Archiving Methods and Approaches: A Comparative Study. *Library Trends*, 54(1), 72–90. Letöltés helye: <https://search.proquest.com/docview/220467286?accountid=27464>
- Massis, B. (2016). Libraries and digital memory. *New Library World*, 117(9/10), 673–676. Letöltés helye: <https://search.proquest.com/docview/1830312026?accountid=27464>
- Matúšková, J., & Hausleitner, Ing. P. (2019). Archiving of Websites and Born Digital Documents in Slovakia. In *8th Colloquium Library and Information Experts of the V4+*

- Countries. Letöltés: 2020. január 14; helye: https://kolokvium.ulib.sk/wp-content/uploads/2019/06/B32_P_Matuskova_SK.pdf
- Mayr, M., & Predikaka, A. (2016). Nationale Grenzen im World Wide Web – Erfahrungen bei der Webarchivierung in der Österreichischen Nationalbibliothek. *Bibliothek Forschung Und Praxis*, 40(1), 90–95. doi: 10.1515/bfp-2016-0007. Letöltés helye: <https://doi.org/10.1515/bfp-2016-0007>
- Michel, A. (2019, július 6). The legal framework for web archiving: Focussing on GDPR [Presentation]. Letöltés 2020. augusztus 20; helye: <https://digital.library.unt.edu/ark:/67531/metadc1608995/>
- Microsoft Silverlight. (2020). Letöltés 2020. augusztus 15; helye: <https://www.microsoft.com/silverlight/>
- Mobile operating system. (2020). In *Wikipedia*. Letöltés helye: https://en.wikipedia.org/w/index.php?title=Mobile_operating_system&oldid=973051284
- National Archives. (1996). UK Government Web Archive. Letöltés: 2020. Január 15; helye: <http://www.nationalarchives.gov.uk/webarchive/>
- National Library of Australia. (2019). TROVE DATABASE. Letöltés ideje: 2020. január 14; helye: <https://trove.nla.gov.au/general/about>
- Németh, M. (2017). Nemzetközi körkép a webarchiválás gyakorlatáról. *Könyvtári figyelő*, 63(4), 575–582. Letöltés helye: https://epa.oszk.hu/00100/00143/00349/pdf/EPA00143_konyvtari_figyelo_2017_04_575-582.pdf
- Németh, M. (2018a). A webarchiválásról történeti megközelítésben. *Könyv, könyvtár, könyvtáros*, 27(2), 48–52. Letöltés helye: <http://ki2.oszk.hu/3k/2018/06/a-webarchivalasrol-torteneti-megkozelitesben/>
- Németh, M. (2018b). *How to catalogue a web archive?* Előadás Information Interactions 2018 workshop, Bratislava. Letöltés ideje 2020. január 8; helye: http://mekosztaly.oszk.hu/mia/doc/How_to_catalogue_a_web_archive.pptx
- Németh, M. (2018c). *The education of web archiving*. Előadás CDA 2018 conference, Bratislava. Letöltés ideje 2020. január 8; helye: http://mekosztaly.oszk.hu/mia/doc/The_education_of_web-archiving_CDA_2018.pptx
- Németh, M. (2019a). *Metadata supported full text search in a web archive*. Előadás IIPC conference, Zagreb. Letöltés ideje 2020. január 8; helye: http://mekosztaly.oszk.hu/mia/doc/Metadata_supported_full-text_search_in_a_web_archive-IIPC-2019-Zagreb.pptx
- Németh, M. (2019b). Nemzetközi tanácskozás az internet megőrzéséről. Letöltés 2020. augusztus 17; helye: https://nemzetikonyvtar.blog.hu/2019/06/18/nemzetkozi_tanacskozaz_az_internet_mego_rzeserol
- Németh, M. (2019c). Using semantic microformats for web archiving – an initial project conception. In Katarina Tomková (Szerk.), *Nové trendy a východiská pri budovani LTP archivov: Zborník príspevkov zo 4. Medzinárodnej konferencie o dlhodobej archivácii Bratislava, 5. 11. 2019*. (pp. 31–38). Bratislava: Univerzitná knižnica v Bratislave. Letöltés ideje 2020. január 14; helye: <https://cloud.ulib.sk/index.php/s/JITVgF9TmRruFtC>
- Németh, M. (2019d). Web Archives as a Research Subject. In *Information and technology transforming lives: Connection, interaction, innovation proceedings* (pp. 471–478). Osijek: Faculty of humanities and social sciences, University of Osijek. Letöltés helye: http://bobcatsss2019.ffos.hr/docs/bobcatsss_proceedings.pdf
- Németh M. (2019e). Webarchiválás két szakmai rendezvény tükrében. *Könyv, könyvtár, könyvtáros*, 28(6), 26–29.
- Németh M. (2020a). Az OSZK Webarchívum új honlapjának felépítése és szolgáltatásai. *Könyv, könyvtár, könyvtáros*, 29.(6.), 16–26. Letöltés helye: https://epa.oszk.hu/01300/01367/00329/pdf/EPA01367_3K_2020_06_016-026.pdf

- Németh, M. (2020b). Exploring special web archives collections related to COVID-19: The case of the National Széchényi Library in Hungary [E-paper]. Letöltés helye: https://cc.au.dk/fileadmin/user_upload/WARCnet/Geeraert_et_al_COVID-19_Hungary.pdf
- Németh, M. (2020). Webarchívum mint a tudományos kutatások tárgya. *Tudományos és Műszaki Tájékoztató*, 67(12), 757–765. Letöltés helye: <https://tmt.omikk.bme.hu/tmt/article/view/12804/14543>
- Németh, M., & Drótos, L. (2019a). A blended learning-based curriculum on Web archiving in the National Széchényi Library. *Digital Library Perspectives*, 35(2), 97–114. doi: 10.1108/DLP-03-2019-0012. Letöltés helye: <https://doi.org/10.1108/DLP-03-2019-0012>
- Németh, M., & Drótos, L. (2019b). Metadata Management and Future Plans to Generate Linked Open Data in the Hungarian Web Archiving Pilot Project. *ITLIB*, 2019(2). Letöltés helye: <https://itlib.evtisr.sk/buxus/docs/38-metadata.pdf>
- Németh, M., & Drótos, L. (2020). Egyedi mentésekre szolgáló webarchiváló szoftverek. *Könyv, könyvtár, könyvtáros*, 29(12), 3–11. Letöltés helye: https://epa.oszk.hu/01300/01367/00334/pdf/EPA01367_3K_2020_12_003-011.pdf
- NetArchiveSuite. MIA WIKI (2017). Letöltés 2019. február 23; helye: <http://mekosztaly.oszk.hu/mediawiki/index.php/NetarchiveSuite>
- NetArchiveSuite documentation. (2018). Letöltés 2020. február 23; helye: <https://sbforge.org/display/NAS/NetarchiveSuite>
- Netarkivet. (2015a). Indsamling materialer til Netarkivet. Letöltés 2020. január 8; helye: http://netarkivet.dk/wp-content/uploads/2015/10/Politik_for_indsamling_af_materiale_til_Netarkivet.pdf
- Netarkivet. (2015b). Pligtafløring til webstedejere. Letöltés 2020. január 8; helye: <http://netarkivet.dk/til-webstedejere/pligtafløring/>.
- Netarkivet. (2018). Solrwayback. Letöltés 2020. február 23; helye: <https://github.com/netarchivesuite/solrwayback#about-solrwayback>
- Netarkivet. (2020). Netarkivet | Adgang til Netarkivet. Letöltés 2020. augusztus 17; helye: <http://netarkivet.dk/adgang/>
- Netlab. (2018a). NETLAB Web Archiving Course Brochure. Letöltés 2019. január 28; helye: <http://www.netlab.dk/wp-content/uploads/2017/04/NetLab-Web-Archiving-Course-Brochure.pdf>
- Netlab. (2018b). Netlab-courses. Letöltés 2019. január 28; helye: <http://www.netlab.dk/services/courses/>
- Nielsen, J. (2016). *Using web archives in research – an introduction* (1.). Aarhus: Netlab. Letöltés helye: http://www.netlab.dk/wp-content/uploads/2016/10/Nielsen_Using_Web_Archives_in_Research.pdf
- Nimbus screen capture. (2018). Letöltés 2020. február 23; helye: http://mekosztaly.oszk.hu/mediawiki/index.php/Nimbus_Screen_Capture
- Niu, J. (2012). An Overview of Web Archiving. *D-Lib Magazine*, 18(3–4). doi: 10.1045/march2012-niu1. Letöltés helye: <https://doi.org/10.1045/march2012-niu1>
- Nuys, C. van, Albertsen, K., Pedersen, L., & Szalóki, G. (2005). A Paradigma projekt. *Tudományos és műszaki tájékoztató*, 52(11–12). Letöltés helye: <https://tmt.omikk.bme.hu/tmt/article/view/1372/1838>
- Országos Széchényi Könyvtár. (2018). Pywb search interface -NSZL. Letöltés 2020. február 23; helye: <http://webadmin.oszk.hu/pywb/>
- Osborne, A. (2020). Heritrix wiki. Letöltés 2020. február 23; helye: <https://github.com/internetarchive/heritrix3/wiki>
- OSZK. (2007). Bibliotheca Nationalis Hungariae, „Az Országos Széchényi Könyvtár programja, 2007. Letöltés helye: http://www.oszk.hu/sites/default/files/szakmai_munkaterv_2007_0.pdf
- OSZK Webarchiválási munkacsoport. (2017). *Az OSZK webarátás pilot projektjének gyűjtőköri tervezete*. Letöltés: 2020. augusztus 17; helye:

- https://webarchivum.oszk.hu/wp-content/uploads/2020/02/Webarchivum_gyujtokor_tervezet.pdf
- OSZK Webarchiválási munkacsoport. (2018). *Részletes beszámoló a 2017-2018 évi webarchiválási pilot projektről*. Letöltés: 2020. augusztus 17; helye: https://webarchivum.oszk.hu/wp-content/uploads/2020/02/Webarchivalas_vezetoi_osszefoglalo_2017-2018.pdf
- Pandora Archive—NLA Selection Guidelines. (2009). Letöltés 2020. augusztus 17; helye: <http://pandora.nla.gov.au/selectionguidelines.html>
- Pandora Archive—Preserving and Accessing Networked Documentary Resources of Australia. (1999, május). Letöltés 2020. augusztus 17; helye: <http://pandora.nla.gov.au/>
- Paterson, J. (2015). *Snapshot of Web harvesting in Canada*. Canada, North America: SelectedWorks. Letöltés helye: <http://search.ebscohost.com/login.aspx?authype=ip,cookie,cpid&custid=s6213251&roupid=main&profile=eds>
- Paul-Choudhury, S. (2011). Your digital legacy. *New Scientist*, 210(2809), 40–43. Letöltés helye: <https://www.newscientist.com/article/mg21028091-400-digital-legacy-the-fate-of-your-online-soul/>
- Peternák, M. (2017). 404 és akiknek nem kell—Médiamúzeum Magyarországon: A „Magyar Tartalom” húsz éves jubileuma alkalmából. Budapest: c3 kulturális Központ. Letöltés: 2020. február 14; helye: http://mekosztaly.oszk.hu/mia/doc/workshop/Peternak_Miklos_404_OSZK_2017_10_13_ea.pptx
- Phillips, M. E., & Phillips, K. K. (2018). End of Term 2016 Presidential Web Archive. *Against the Grain*, 29(6), 27. Letöltés helye: <https://digital.library.unt.edu/ark:/67531/metadc1114874/>
- Puppeteer. (2020). Letöltés 2020. február 23; helye: <https://pptr.dev/>
- Pywb. (2019). Letöltés 2020. február 23; helye: <http://mekosztaly.oszk.hu/mediawiki/index.php/PyWb>
- RealAudio. (2020). In *Wikipedia*. Letöltés helye: <https://en.wikipedia.org/w/index.php?title=RealAudio&oldid=972363209>
- Rich web application. (2020). In *Wikipedia*. Letöltés helye: https://en.wikipedia.org/w/index.php?title=Rich_web_application&oldid=970947425
- Riigikogu. Legal Deposit Copy Act—Estonia (2016). Estonia. Letöltés helye: <https://www.riigiteataja.ee/en/eli/514092016001/consolide>
- Riva P., Boeuf P. L., Žumer M., & Szerkesztőbizottsága E. (2017). IFLA könyvtári referenciamodel, 99. Letöltés helye: https://www.ifla.org/files/assets/cataloguing/frbr-irm/ifla-irm-august-2017_rev201712-hu.pdf
- Robots exclusion standard—Wikipedia. (2020). Letöltés 2020. augusztus 14; helye: https://en.wikipedia.org/wiki/Robots_exclusion_standard
- Rosenthal, David. S. (2017). Losing the Battle to Archive the Web. Letöltés 2020. január 7; helye: <http://dponline.org/blog/idpd/losing-the-battle-to-archive-the-web-%20>
- Ruusalep, R. (2017). Preserving digital legal deposit—New challenges and opportunities. In *IFLA WLIC 2017 – Wrocław, Poland – Libraries. Solidarity. Society. In Session 210—Preservation and Conservation (PAC) Strategic Programme*. Wrocław: IFLA -- International Federation of Library Associations and Institutions. Letöltés helye: <http://library.ifla.org/id/eprint/1677>
- Schostag, S. (2019a). Harvesting YouTube videos—A manual approach [Presentation]. Letöltés 2020. augusztus 20; helye: <https://digital.library.unt.edu/ark:/67531/metadc1609005/>
- Schostag, S. (2019b, július 6). Why, what, when and how?: Curatorial decisions and the importance of their documentation [Presentation]. Letöltés 2020. augusztus 20; helye: <https://digital.library.unt.edu/ark:/67531/metadc1608959/>
- Schostag, S., & Fonss-Jorgensen, E. (2012). Webarchiving: Legal Deposit of Internet in Denmark. A Curatorial Perspective. *Microform & Digitization Review*, 41(3–4), 110–

120. Letöltés helye:
<https://search.proquest.com/docview/1520327503?accountid=27464>
- Sierman, B., & Teszelszky, K. (2017). How can we improve our web collection? An evaluation of webarchiving at the KB National Library of the Netherlands (2007-2017), 27(2), 94–106. doi: 10.1177/0955749017725930. Letöltés helye:
<https://doi.org/10.1177/0955749017725930>
- Signori, B. (2017). Preserving cultural heritage: Better together! In *IFLA WLIC 2017 – Wrocław, Poland – Libraries. Solidarity. Society. In Session S08—Satellite Meeting: Preservation and Conservation Section joint with the Association International Francophone des Bibliothécaires et Documentalistes (AIFBD) in collaborati*. Wrocław: IFLA -- International Federation of Library Associations and Institutions. Letöltés helye: <http://library.ifla.org/id/eprint/1801>
- SolrMIA. (2018). Letöltés 2020. február 23; helye: <http://193.6.201.202/solrmia/>
- Solrwayback kereső. (2018). Letöltés 2020. február 23; helye:
<http://webadmin.oszk.hu/solrwayback/>
- Stirling, P., Illien, G., Sanz and, P., & Sepetjan, S. (2012). The state of e-legal deposit in France: Looking back at five years of putting new legislation into practice and envisioning the future. *IFLA Journal*, 38(1), 5–24. Letöltés helye:
<http://10.0.4.153/0340035211435323>
- Storarr, T., & Seaman, G. (2016). Researching User Needs: The UK Government Web Archive—UNT Digital Library. Letöltés 2020. augusztus 20; helye:
<https://digital.library.unt.edu/ark:/67531/metadc1477155/>
- Strodl, S., Becker, C., Neumayer, R., & Rauber, A. (2007). How to Choose a Digital Preservation Strategy: Evaluating a Preservation Planning Procedure. In *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 29–38). New York, NY, USA: ACM. doi: 10.1145/1255175.1255181 Letöltés helye:
<https://doi.org/10.1145/1255175.1255181>
- Sylvie-Rollason-Cass. (2020). What is brozzler? Letöltés 2020. február 23; helye:
<https://support.archive-it.org/hc/en-us/articles/360000343186-What-is-Brozzler->
- Szabo, S. (2018, november 15). Website Archiving Guide Dutch National Archives [Presentation]. Letöltés 2020. augusztus 20; helye:
<https://digital.library.unt.edu/ark:/67531/metadc1477176/>
- SZTAKI. (2011). Living Web Archive—LIWA. Letöltés 2020. augusztus 18; helye:
<https://www.sztaki.hu/en/projects/liwa>
- SZTAKI. (2013). LAWA - Longitudinal Analytics of Web Archive Data. Letöltés 2020. augusztus 18; helye:
https://old.sztaki.hu/tudomany/projektek/projekt_informaciok/?uid=00303
- Temporal drift – MIA. (2017). Letöltés 2020. augusztus 18., helye:
https://webarchivum.oszk.hu/mediawiki/index.php?title=Temporal_drift
- Teszelszky, K. (2017). The harvest of the Dutch digital fields: The landscape of webarchiving in The Netherlands. Letöltés ideje: 2020. január 8; helye:
https://webarchivum.oszk.hu/wp-content/uploads/2020/03/Kees_Teszelszky_2017_Presentatie_webarchivering_KB_BUDAPEST_404.pptx
- Teszelszky, K. (2019a). Distant reading: The Frisian Web Domain. Letöltés 2020. január 8; helye:
http://mekosztaly.oszk.hu/mia/doc/DH_2019/2019.09.11_Teszelszky_Friese_web.pptx
- Teszelszky, K. (2019b). Web archaeology in The Netherlands: The selection and harvest of the Dutch web incunables of provider Euronet (1994–2000). *Internet Histories*, 3(2), 180–194. doi: 10.1080/24701475.2019.1603951. Letöltés helye:
<https://doi.org/10.1080/24701475.2019.1603951>
- The dead web – The end | online kiállítás | Ludwig Múzeum. (2020). Letöltés 2020. augusztus 17., helye: <https://www.ludwigmuseum.hu/valogatás/dead-web-end-online-kiállítás>

- Thomson, S. D., & Kilbride, W. (2015). Preserving Social Media: The Problem of Access. *New Review of Information Networking*, 20(1/2), 261–275. Letöltés helye: <http://10.0.4.56/13614576.2015.1114842>
- Thornton, J. B. (2012). Archiving of Comprehensive Annual Financial Reports (CAFRs) on State Government Web Sites. *Behavioral & Social Sciences Librarian*, 31(2), 87–95. doi: 10.1080/01639269.2012.686244. Letöltés helye: <http://dx.doi.org/10.1080/01639269.2012.686244>
- Tiiu, D. (2018, november 13). Becoming a Web Archivist: My 10 Year Journey in the National Library of Estonia [Presentation]. Letöltés 2020. augusztus 20; helye: <https://digital.library.unt.edu/ark:/67531/metadc1477143/>
- Togan, M., & Florea, I. (2015). A Reference Model for a Trusted Service Guaranteeing Web-content. In Reimer, H., Pohlmann, N., & Schneider, W. (Szerk.), *ISSE 2015* (pp. 216–224). Wiesbaden: Springer Fachmedien Wiesbaden. doi: 10.1007/978-3-658-10934-9_18. Letöltés helye: https://doi.org/10.1007/978-3-658-10934-9_18
- UNESCO. Charter on the Preservation of Digital Heritage (2003). Letöltés ideje 2020. augusztus 18; helye: http://portal.unesco.org/en/ev.php-URL_ID=17721&URL_DO=DO_TOPIC&URL_SECTION=201.html
- Veikkolainen, P., Lager, L., Lehtonen, J., & Kaunonen, K. (2016, december 4). Long-Term Preservation of the Finnish Web Archive [Presentation]. Letöltés 2020. augusztus 20; helye: <https://digital.library.unt.edu/ark:/67531/metadc1477130/>
- Visky, Á. L. (2020). Gyorsmérleg – az OSZK Webarchívumának és néhány könyvtárnak a KDS-K pályázat keretében történt együttműködéséről. *Könyv, könyvtár, könyvtáros*, 29.(7-8.), 24-38. Letöltés helye: http://epa.oszk.hu/01300/01367/00330/pdf/EPA01367_3K_2020_07-08_024-038.pdf
- Visky, Á. L., & Drótos, L. (2020). Rákóczi-archívum. *Könyv, könyvtár, könyvtáros*, 29(3), 35–48. Letöltés helye: https://epa.oszk.hu/01300/01367/00326/pdf/EPA01367_3K_2020_03_035-048.pdf
- Vitéz, G. (2020). Az arató és a megjelenítő szoftverekről. Letöltés 2020. február 23; helye: <http://mekosztaly.oszk.hu/mia/e-learning/prezentaciok/VG-eloadas.pptx>
- Vozár, Z. (2019). CZ web archive development. Data processing and provision in cluster. In *8th Colloquium Library and Information Experts of the V4+ Countries*. Letöltés ideje: 2020. február 15; helye: https://kolokvium.ulib.sk/wp-content/uploads/2019/06/B34_P_Vozar_CZ.pdf
- VRML Virtual Reality Modeling Language. (1995). Letöltés 2020. augusztus 15; helye: <https://www.w3.org/MarkUp/VRML/>
- W3C Consortium. (2019). Introduction to Web Accessibility. Letöltés 2020. augusztus 17; helye: <https://www.w3.org/WAI/fundamentals/accessibility-intro/>
- WAIL - MIA WIKI. (2017). Letöltés 2020. február 23; helye: <http://mekosztaly.oszk.hu/mediawiki/index.php/WAIL>
- Wappalyzer—Identify technology of websites. (2020). Letöltés 2020. augusztus 15; helye: <https://www.wappalyzer.com/>
- WARC – MIA. (2017). Letöltés 2020. augusztus 17; helye: <https://webarchivum.oszk.hu/mediawiki/index.php?title=WARC>
- WCT manuals. (2019). Letöltés 2020. február 23; helye: <http://webcurator.sourceforge.net/manuals.shtml>
- WEB Archive UK & JICS. (2015). Shine Project Historical Research Prototype. Letöltés ideje 2020. február 14; helye: <https://www.webarchive.org.uk/shine>
- Web Curator Tool. (2018). Letöltés 2020. február 23; helye: <http://mekosztaly.oszk.hu/mediawiki/index.php/WCT>
- Webrecorder—MIA WIKI. (2018). Letöltés 2020. február 23; helye: <http://mekosztaly.oszk.hu/mediawiki/index.php/Webrecorder>
- Webrecorder/warcit*. (2020). Python, Webrecorder. Letöltés helye: <https://github.com/webrecorder/warcit>
- Website Archiving and Monitoring Solutions | MirrorWeb. (2020). Letöltés 2020. augusztus 20., helye: <https://www.mirrorweb.com>

- Wichtowski, R. (2017). Increasing Copyright Protection for Social Media Users by Expanding Social Media Platforms' Rights. *Duke Law & Technology Review*, 16(1), 253–268. Letöltés helye: <https://scholarship.law.duke.edu/dltr/vol15/iss1/13/>
- Winters, J. (2015, április 27). Big UK Domain Data for the Arts and Humanities [Presentation]. Letöltés 2020. augusztus 20; helye: <https://digital.library.unt.edu/ark:/67531/metadc1476406/>
- Winters, J. (2017). Breaking in to the mainstream: Demonstrating the value of internet (and web) histories. *Internet Histories*, 1(1–2), 173–179. doi: 10.1080/24701475.2017.1305713. Letöltés helye: <https://doi.org/10.1080/24701475.2017.1305713>
- XHTML+RDFa 1.1—Third Edition. (2015). Letöltés 2020. augusztus 14 helye: <https://www.w3.org/TR/xhtml-rdfa/>
- Zeng, M., & Quin, J. (2016). *Metadata*. Facet Publishing.



Nyilvántartási szám: DEENK/62/2021.PL
Tárgy: PhD Publikációs Lista

Jelölt: Németh Márton
Doktori Iskola: Informatikai Tudományok Doktori Iskola
MTMT azonosító: 10049742

A PhD értekezés alapjául szolgáló közlemények

Magyar nyelvű tudományos közlemények hazai folyóiratban (6)

1. **Németh, M.:** Az OSZK Webarchívum új honlapjának felépítése és szolgáltatásai.
Kv. kv. kv. 29 (6), 16-26, 2020. ISSN: 1216-6804.
2. Drótos, L., **Németh, M.:** Egyedi mentésekre szolgáló webarchiváló szoftverek.
Kv. kv. kv. 29 (12), 3-11, 2020. ISSN: 1216-6804.
3. **Németh, M.:** Webarchívum mint a tudományos kutatások tárgya.
Tud. műsz. táj. 67 (12), 757-765, 2020. ISSN: 0041-3917.
4. **Németh, M.:** A digitális könyvtárak, a digitális könyvtáros és a jövő perspektívái.
Kv. kv. kv. 27 (8), 7-17, 2018. ISSN: 1216-6804.
5. Horváth, Á., **Németh, M.:** Kulturális közkinccsek feltárása nyílt bibliográfiai adatok révén.
Kvt. figy. 63 (2), 196-210, 2017. ISSN: 0023-3773.
6. **Németh, M.:** Nemzetközi körkép a webarchiválás gyakorlatáról.
Kvt. figy. 63 (4), 575-582, 2017. ISSN: 0023-3773.

Idegen nyelvű tudományos közlemények külföldi folyóiratban (3)

7. **Németh, M., Drótos, L.:** A blended learning-based curriculum on Web archiving in the national Széchényi library.
DLP. 35 (2), 97-114, 2019. ISSN: 2059-5816.
DOI: <http://dx.doi.org/10.1108/DLP-03-2019-0012>
8. **Németh, M., Drótos, L.:** Metadata Management and Future Plans to Generate Linked Open Data in the Hungarian Web Archiving Pilot Project.
ITLib. 2, 38-44, 2019. ISSN: 1335-793X.
9. **Németh, M., Simon, A.:** Public collections on the Semantic web in a Hungarian context.
ITLib. 2, 66-74, 2016. ISSN: 1335-793X.





Egyéb folyóiratközlemények (7)

10. **Németh, M.**: A könyvtári szemantikus webes fejlesztések világa: gondolatok az SWIB konferenciasorozat kapcsán.
Tud. műsz. táj. 66 (3), 157-160, 2019. ISSN: 0041-3917.
11. **Németh, M.**: Webarchiválás két szakmai rendezvény tükrében.
Kv. kvt. kvt. 28 (26), 29, 2019. ISSN: 1216-6804.
12. Dancs, S., Koltay, T., **Németh, M.**: Adat- és tudásmenedzsment a posztgazdaság korában: mozaikok az INFINT Information Interactions 2018 nemzetközi konferenciáról.
Kv. kvt. kvt. 27 (11), 39-43, 2018. ISSN: 1216-6804.
13. **Németh, M.**: A webarchiválás nemzetközi környezete: mozaikok az IIPC 2019 kongresszusról.
Kv. kvt. kvt. 27 (12), 23-27, 2018. ISSN: 1216-6804.
14. **Németh, M.**: A webarchiválásról történeti megközelítésben: [Ismeretítés].
Kv. kvt. kvt. 27 (2), 48-52, 2018. ISSN: 1216-6804.
15. **Németh, M.**: Hogyan tudjuk fejleszteni a webes gyűjteményünket?: a Holland Nemzeti Könyvtár webarchiválási tevékenységének értékelése (2007-2017) : [Ismeretítés].
Kvt. figy. 64 (4), 634, 2018. ISSN: 0023-3773.
16. **Németh, M.**: "404 Not Found - Ki őrzi meg az internetet?": webarchiválás workshop az Országos Széchényi Könyvtárban.
Tud. műsz. táj. 64 (11), 577-582, 2017. ISSN: 0041-3917.

Magyar nyelvű konferencia közlemények (3)

17. **Németh, M.**: A mikroadatok felhasználása webarchiválási környezetben.
In: Networkshop 2020 Online : Országos Online Konferencia : 2020. szeptember 2-4.. Szerk.: Tick József, Kokas Károly, Holl András, Hungarnet, Budapest, 81-87, 2020. ISBN: 9786150103761
18. Drótos, L., **Németh, M.**: A webarchiválás oktatása.
In: Networkshop 2018 konferenciakötet. Szerk.: Tick József, Kokas Károly, Holl András, Hungarnet, Budapest, 31-37, 2018.
19. Drótos, L., **Németh, M.**: Az OSZK-ban folyó kísérleti webarchiválási projekt első évének tapasztalatai.
Tud. műsz. táj. 65 (7-8), 389-400, 2018. ISSN: 0041-3917.

Idegen nyelvű konferencia közlemények (5)

20. **Németh, M.**: Using semantic microformats for web archiving: an initial project conception.
In: CDA 2019 Konferenciakötet, [S.n.], Pozsony, 31-38, 2019.





21. **Németh, M.**, Drótos, L.: Web Archives as a Research Subject.
In: Information and technology transforming lives: connection, interaction, innovation :
Proceedings of the XXVII. BOBCATSS Symposium. Ed.: Gordana Gašo, Mirna Gilman
Ranogajec, Jure Žilić, Madeleine Lundman, University of Osijek, Osijek, 471-478, 2019.
ISBN: 9789533141213
22. Drótos, L., **Németh, M.**: The education of web archiving.
In: CDA 2018 Trvalá udržateľnosť a perspektívy ďalšieho rozvoja LTP archívov. Ed.:
Tomková Katarína, Univerzitná knižnica v Bratislava, Bratislava, 108-116, 2018.
23. **Németh, M.**, Drótos, L.: Web-Archiving: responsibility of Public Collections in the Preservation of
the Digital Culture.
In: Bobcatss 2018 Riga : The power of reading : proceedings. Eds.: Lelde Petrovska, Baiba
Ivane-Kronberga, Zane Meldere, University of Latvia Press, Riga, 124-126, 2018. ISBN:
9789934183539
24. **Németh, M.**: Semantic web developments in Hungarian public collections in an international
perspective.
Libellarium. 9 (2), 169-186, 2016. ISSN: 1846-8527.
DOI: <http://dx.doi.org/10.15291/libellarium.v9i2.262>

Idegen nyelvű absztrakt kiadványok (2)

25. **Németh, M.**, Drótos, L.: How to catalogue a web archive? Some solutions for metadata
management at the web harvesting pilot project of National Széchényi Library, Hungary.
In: Information Interactions 2018 : International conference on digital environment and
information science : Book of Abstracts. Ed.: Jela Steinerová, Katarína Buzová, Comenius
University in Bratislava, Bratislava, 10, 2018. ISBN: 9788022346047
26. **Németh, M.**, Drótos, L.: Hungarian web archiving pilot project in the National Széchényi Library.
In: 2017 8TH IEEE International Conference on Cognitive Infocommunications : CogInfoCom,
Institute of Electrical and Electronics Engineers Inc, Piscataway, 209-211, 2017,
(International Conference on Cognitive Infocommunications, ISSN 2375-1312) ISBN:
9781538612644

További közlemények



Idegen nyelvű, külföldi könyvrészletek (2)

27. **Németh, M.**: Folk culture, literature and nation building efforts in the Romantic Age of Central
Europe.
In: Integrity in integration : Developing sustainable dialogue / Rachael Weber, Nagypál
Szabolcs, Sajda Peter, Oikumené Akadémická YMCA, Prága, 44-50, 2008, (Ecumenical
Anthology ; 7.) ISBN: 978963869469



28. **Németh, M.**: Nationalism, Xenophobia, Populism and Nation States in the Modern History of Central Europe.
In: Roots of Vision, Routes of Change Nationalism and Ecumenical Leadership. Eds.: Rachael Weber, Peter Sajda, Nagypál Szabolcs, BGÖI & WSCF-CESR, Budapest-Bratislava, 34-43, 2007. ISBN: 9789638694652
- Magyar nyelvű tudományos közlemények hazai folyóiratban (17)
29. **Németh, M.**: De hát hol a könyvtár mely célhoz vezet?: az északi országok könyvtárügye a társadalmi mobilitás és kohézió szolgálatában.
Észak. 3 (1), 7-20, 2020. ISSN: 2560-2519.
30. Békésiné Bognár, N. E., **Németh, M.**: Hogyan lépünk egyről a kettőre?: a K2 műhely múltja és jelene.
Kvt. figy. 30 (4), 624-636, 2020. ISSN: 0023-3773.
31. **Németh, M.**: Nemzeti nyelvű könyvtári e-könyvszolgáltatási keretrendszerek: skandináv esettanulmányok itthoni tanulságokkal.
Tud. műsz. táj. 64 (5), 247-251, 2017. ISSN: 0041-3917.
32. **Németh, M.**: Online oktatási tartalomszolgáltatási projekt Nagy Britanniában: a BBC, az oktatás és a kulturális, közgyűjteményi szféra közös platformon.
Kv. kvt. kvt. 26 (2), 30-32, 2017. ISSN: 1216-6804.
33. Bánki, Z., Mészáros, T., **Németh, M.**, Simon, A.: Azonos személyekre vonatkozó név besorolási rekordok automatikus felderítése a PIM adatbázisában.
Tud. műsz. táj. 63 (12), 471-477, 2016. ISSN: 0041-3917.
34. Berger, Á., **Németh, M.**, Simon, A.: Az Europeana Inside és magyar megvalósítása.
Tud. műsz. táj. 61 (9), 342-349, 2014. ISSN: 0041-3917.
35. **Németh, M.**: A Digital Library Learning nemzetközi mesterképzésről - hallgatói szemmel.
Kv. kvt. kvt. 22 (9), 46-52, 2013. ISSN: 1216-6804.
36. **Németh, M.**: Bibzoom - a dán könyvtárak kulturális, zenei szolgáltató-ismeretterjesztő portálja.
Kv. kvt. kvt. 22 (12), 42-45, 2013. ISSN: 1216-6804.
37. Horváth, S. D., **Németh, M.**: Bibliotheca pro futuro.
Kv. kvt. kvt. 20 (10), 21-24, 2011. ISSN: 1216-6804.
38. **Németh, M.**: Közönségkapcsolatok, könyvtári márkanév, civil munka és a szakmai tartalom egyensúlya az észak-európai országokban.
Kv. kvt. kvt. 20 (7), 39-42, 2011. ISSN: 1216-6804.
39. **Németh, M.**: Stratégiai tervezési körkép az észak-európai könyvtárügyről.
Kv. kvt. kvt. 19 (10), 35-42, 2010. ISSN: 1216-6804.
40. **Németh, M.**: Trendek és szerzői jogi problémák a digitális médiaszolgáltatás világában.
Kvt. figy. 56 (2), 263-268, 2010. ISSN: 0023-3773.





41. **Németh, M.:** Áttekintés a modern online audiovizuális szolgáltatások világáról.
Kvt. figy. 55 (1), 29-36, 2009. ISSN: 0023-3773.
42. **Németh, M.:** Az észak-európai országok közönyvtárügyéről.
Kv. kvt. kvt. 18 (1), 48-55, 2009. ISSN: 1216-6804.
43. **Németh, M.:** Az Europeana és a TELPlus projekt.
Kv. kvt. kvt. 18 (3), 3-7, 2009. ISSN: 1216-6804.
44. **Németh, M.:** Az e-könyvek szolgáltatása profitorientált dokumentumticipi körülmények között.
Kvt. figy. 54 (2), 234-240, 2008. ISSN: 0023-3773.
45. Horváth, Á., **Németh, M.:** Az OSZK új közösségi világháló (web 2.0) fejlesztései.
Kv. kvt. kvt. 17 (8), 34-37, 2008. ISSN: 1216-6804.

Idegen nyelvű tudományos közlemények külföldi folyóiratban (2)

46. Bánki, Z., Mészáros, T., **Németh, M.**, Simon, A.: Checking the identity of entities by machine algorithms: the next step to the Hungarian National Namespace.
Code4Lib. 33, 11765, 2016. ISSN: 1940-5758.
47. **Németh, M.:** Hungarian Libraries and Librarianship, 1990-2013: an Overview.
Library Trends. 63 (2), 212-232, 2014. EISSN: 1559-0682.
DOI: <http://dx.doi.org/10.1353/lib.2014.0034>

Egyéb folyóiratközlemények (27)

48. **Németh, M.:** Információs műveltség és a könyvtárak a hamis hírek korszakában.
Kvt. figy. 66 (3), 507-510, 2020. ISSN: 0023-3773.
49. **Németh, M.:** Szubjektív mozaikok a Networkshop 2020 konferenciáról.
Kvt. figy. 30 (4), 668-675, 2020. ISSN: 0023-3773.
50. **Németh, M.:** 85. IFLA World Library Information Congress 2019, Athén - rövid szubjektív áttekintés.
Könyvtárvilág. 7 (4), 1, 2019. EISSN: 2064-1443.
51. Ford. **Németh, M.:** Áttekintés a dán könyvtárosok adattudományt érintő kompetenciáiról és gyakorlati tevékenységeiről.
Kvt. figy. 65 (4), 611-615, 2019. ISSN: 0023-3773.
52. Arany, Z., Nagy, A., **Németh, M.**, Radó, R.: Mozaikok az észéki Bobcatss 2019 konferenciáról.
Kv. kvt. kvt. 28 (2), 16-23, 2019. ISSN: 1216-6804.
53. **Németh, M.:** Színes körkép egy összetett könyvtári területről.
Kvt. figy. 65 (4), 616-618, 2019. ISSN: 0023-3773.
54. **Németh, M.:** A REACH projekt: közösségi platform a társadalmi részvétel alapuló tevékenységek, illetve a közösségi innováció elősegítésének céljából.
Kv. kvt. kvt. 27 (5), 18-20, 2018. ISSN: 1216-6804.





55. **Németh, M.**: Digitális Bölcsészet. Gondolatok egy új nyílt hozzáférésű, interdiszciplináris e-folyóirat születése kapcsán.
Kv. kv. kv. 27 (12), 48-52, 2018. ISSN: 1216-6804.
56. **Németh, M., Nagy, A.**: Mozaikok a rigai Bobcatsss 2018 konferenciáról.
Kv. kv. kv. 27 (2), 11-21, 2018. ISSN: 1216-6804.
57. **Németh, M.**: Szubjektív beszámoló az IFLA wroclawi világkonferenciájáról.
Kvt. figy. 64 (2), 273-278, 2018. ISSN: 0023-3773.
58. Arany, Z., Bálint, D., Farkas, C., Kállai, A., Nagy, A., **Németh, M.**, Tóth, V.: A 25. jubileumi BOBCATSSS konferencia az ezer tő országában.
Kv. kv. kv. 26 (2), 38-41, 2017. ISSN: 1216-6804.
59. **Németh, M.**: Áttekintés a Szakkönyvtári Seregszemle 2017 konferenciáról.
Kv. kv. kv. 26 (6), 21-26, 2017. ISSN: 1216-6804.
60. Ford. **Németh, M.**: Beszélgetésre épülő programok svéd és norvég közkönyvtárakban: esettanulmányok a közkönyvtárak integrációjában betöltött szerepéről.
Kvt. figy. 63 (Különszám), 83-89, 2017. ISSN: 0023-3773.
61. Virág, G., Lendvay, M., **Németh, M.**: Mozaikok az ILIDE 2017 konferenciáról.
Kv. kv. kv. 26 (7), 23-28, 2017. ISSN: 1216-6804.
62. Csikász-Nagy, Á., **Németh, M.**, Pataki, F., Tót, B.: Önkéntesként szerzett benyomások 2017 kiemelkedő nemzetközi szakmai eseményéről: IFLA World Library and Information Congress 83rd IFLA General Conference and Assembly 2017. augusztus 19-25. Wroclaw, Lengyelország.
Kv. kv. kv. 26 (11), 11-17, 2017. ISSN: 1216-6804.
63. Arany, Z., Jávorka, B., Nagy, A., **Németh, M.**: Egy hét az oroszlan barlangjában: BOBCATSSS 2016, Lyon.
Kv. kv. kv. 25 (3), 17-22, 2016. ISSN: 1216-6804.
64. **Németh, M.**: Mozaikok a göteborgi könyvvásárról.
Kvt. figy. 62 (2), 245-250, 2016. ISSN: 0023-3773.
65. **Németh, M.**: Mozaikok az információs műveltség európai konferenciájáról (ECIL, 2015).
Kvt. figy. 62 (1), 86-90, 2016. ISSN: 0023-3773.
66. **Németh, M.**: Oslo, Helsinki, Lyon, Brno: konferencia-mozaikok 2014 tavasza és 2016 januárja között.
Kvt. figy. 61 (1), 75-84, 2015. ISSN: 0023-3773.
67. **Németh, M.**: Szemelvények a MoodleMoot 2015 konferenciáról.
Kv. kv. kv. 24 (9), 28-33, 2015. ISSN: 1216-6804.
68. **Németh, M.**: E-könyvekről sokszemközt: Az ELTE BTK Könyvtár- és Információtudományi Intézetének szakmai napja.
Kvt. figy. 60 (4), 486-491, 2014. ISSN: 0023-3773.





69. **Németh, M.**: Kézikönyv a könyvtár- és információtudomány elméleti alapjairól, a digitális könyvtárakról és a digitális könyvtárosok képzéséről.
Kvt. Figy. 60 (1), 117-131, 2014. ISSN: 0023-3773.
70. **Németh, M.**: Napjaink könyvtári világának aktuális kérdései: ketten ugyanarról a könyvről.
Kvt. figy. 59 (3), 564-570, 2013. ISSN: 0023-3773.
71. **Németh, M.**: A WEB 2.0 hatásai a tudásmenedzsmentre.
Tud. műsz. táj. 57 (2), 73-78, 2010. ISSN: 0041-3917.
72. **Németh, M.**: Benyomások a göteborgi IFLA-kongresszuson.
Kv. kvt. kvt. 19 (10), 32-34, 2010. ISSN: 1216-6804.
73. **Németh, M.**: Elektronikus audiovizuális szolgáltatások dán könyvtárakban.
Kv. kvt. kvt. 19 (3), 38-41, 2010. ISSN: 1216-6804.
74. **Németh, M.**: Médiaműveltségről sok szemmel: Prágában.
Kv. kvt. kvt. 18 (9), 49-51, 2009. ISSN: 1216-6804.

Magyar nyelvű konferencia közlemények (1)

75. **Németh, M.**: Könyvtárosképzés az északi térségben: integrációs utak és lehetőségek.
In: II. Masters of Library and Information Science konferencia : konferenciakötet. Szerk.:
Sebestyén György, Habók Lilla, Nemes László, ELTE BTK, Budapest, 60-65, 2011.

Idegen nyelvű konferencia közlemények (3)

76. **Németh, M.**: iSchool community and Scandinavian library schools.
In: Design, innovation, participation : BOBCATSSS 2015 : Proceedings and abstracts. Eds.:
Pavla Minarikova, Lukas Stuhel, Flow, Brno, 266-273, 2015. ISBN: 9788090548077
77. **Németh, M.**: Digital library education on Master level in Borås, Copenhagen, and Oslo Qualitative case study.
In: Bobcatsss 2014 Conference on Library and Information Science : Library (r)evolution:
Promoting sustainable information practices, [S.n.], Göteborg, 1-6, 2014.
78. **Németh, M.**: Serving Library Users by Effective Communication Ways in Loughborough University Library.
In: From Collections to Connections: Turning Libraries "Inside-Out" : The 21st International BOBCATSSS Conference Proceedings. Eds.: Taskin Cakmak T., Dogan G, Hacettepe University, Ankara, 175-178, 2014. ISBN: 9789754913613

A DEENK a Jelölt által az iDEa Tudóstérbe feltöltött adatok bibliográfiai és tudománymetriai ellenőrzését a tudományos adatbázisok és a Journal Citation Reports Impact Factor lista alapján elvégezte.



**Short thesis for the degree of doctor of
philosophy (PhD)**

**The theoretical and practical fundamental
elements of web archiving**

**The first steps of institutional web
archiving in Hungary**

by Márton Németh

Supervisor:

Mária Borbély Eszenyiné



UNIVERSITY OF DEBRECEN

Doctoral School of Informatics

Debrecen, 2021

The theoretical and practical fundamental elements of web archiving

The first steps of institutional web archiving in Hungary

This Phd thesis is focusing on web-archiving issues. By offering a broad overview of theoretical and practical elements of this discipline mainly in a Hungarian context, however in international perspective, the main aim is to facilitate further research and development activities in this field.

The thesis offers an introduction to some basic conceptions and major challenges of web-archiving. A contextual analysis is being offered about web archiving related to the library, archive, and museum fields. A brief description of the international context is also being described.

The core of the thesis has built on a detailed analysis about the initial conception, professional frameworks, and workflows of web-archiving in the National Széchényi Library, Hungary as the first organized project in that field in this country. The description of major workflow elements, the challenges in collection development and metadata management fields are being described together with the international context of the project. A brief outlook to the main elements of the software and hardware infrastructure, an introduction the challenges in long-term digital preservation context and a brief description of the legal framework together with an outlook to the communication conception of the project are also essential elements of this analysis.

A whole chapter is focusing on the educational context of web archiving by international and Hungarian perspectives. An essential part of the thesis is describing the research perceptions in web archiving field by introducing various new sub-disciplines in this context and offering an overview to theoretical and practical research activities in the future related to the use of semantic microformats.

Main hypotheses and research questions

The research serving as the basis of the PhD dissertation is intending to present the theoretical elements of web archiving, and also aims to introduce some ways to translate theoretical frameworks into practice by the framework of a web

archiving project launched in 2017 at the National Széchényi Library. The initial research hypotheses can be formulated according to five main aspects below:

1. The main hypothesis of the research is whether it is possible to make an overview of the legitimacy of web archiving, its theoretical background, field of interpretation, theoretical and practical limitations, public collection frameworks, and some ways of using this field for research purposes. The international background of this field and a broad overview of the related professional literature also seemed to be offered related to the Hungarian project.
2. An important research question arose as an initial hypothesis that in addition to highlighting the international theoretical and practical framework, the system of archiving and service framework of organized web archiving can be created within the National Széchényi Library, by developing appropriate workflows, and offering appropriate financial, technological, legal requirements. If it cannot be done or is it just only partially done, which are the main challenging factors to prevent building-up the whole framework?
3. It is also essential to clarify the essential features of the currently emerging jobs of web librarians and web curators, and the extent to which these activities cover within the framework of libraries, and what belong to other sectors of public collections and within the framework of informatics.
4. Is there a chance that the professional, financial and information technology infrastructure challenges that may arise together with the cumbersome process of creating the legal and regulatory environment, make it impossible to turn a research and development pilot project into a permanent library service?
5. Is it possible to include the topic of web archiving in the framework of accredited librarian training or place as a subject of higher education programs in Hungary?

Some professional elements of the answers to the hypotheses are being explained by integrating them to the dissertation or can be found among the theses summarizing the results of the research, the overall results are being outlined in the final chapter of the dissertation.

Another unique feature of the dissertation is that the implementation of the individual research project meets in some points with the collective efforts that we made with my colleagues at the Department E-Library Services of the National Széchényi Library (currently have become a separate Department of Web Archiving) in order to establish a permanent web archiving service. To differentiate the various tasks in the text of the dissertation, the use of the singular first person always refers to a content element that is related to individual research, analysis, exploration of research results. The imprints of the collective, practical efforts on web archiving are reflected in the use of the plural first person. I am integrating the practical elements of individual research and the results of collective practical activities (if any) into each chapter, and I am offering a general overview about the evaluation of the research in the final chapter.

Evaluation of the research

In this chapter, I first address the results of research on the contents of the hypotheses.

In my opinion, the objectives set out in the first hypothesis were met. I have tried to verify the relevance of web archiving as a scientific topic from several approaches. I have taken special care to outline its basic public collection framework. In several chapters, the dilemmas related to professional, technological, and theoretical, methodological limitations of the web archiving activities, and the range of interpretation of the concept of web archiving are being discussed. In this context, a set of challenges related to the quality control and long-term preservation of archived files are also being mentioned as a separate topic. Throughout my research in each chapter I wanted to take special focus to the wide-ranging exploration of the relevant international professional literature regarding the practical tasks and theoretical research background of web archiving in Hungary. An important component of the dissertation is an overview of current scientific research areas related to web archiving. Web archiving as a comprehensive interdisciplinary field of professional research is increasingly institutionalized in international academic life. The same cannot be said in Hungarian context. Since web archiving as a permanent professional service will be included in the basic activities of the National Széchényi Library through the

amendment of the Cultural Law, in my hope the emerging collection, as the subject of scientific research, will gain more widespread use in the future. With this dissertation, I wanted to give the initial stimulus to the unfolding of this process. On the one hand, it is necessary to provide a general overview of the general scientific aspects of web archiving. In addition, the main goal of this dissertation is to provide an overview of the various research and development activities related to web archiving. Furthermore, to set the research topics at the forefront of my personal interest, I also want to inspire as many people as possible to choose web archiving and the web archive as a collection as the subject of their scientific research.

In addition to the theoretical and methodological frameworks, the other very important component of the dissertation is the outline of the elements of the emerging Hungarian practice through the contents of the second hypothesis. After discussing the most useful aspects of international practice in Hungary, it was necessary to present the practical workflows that describe the activity. Through a variety of analyses, I answer the question posed in the hypotheses that together with my colleagues we have successfully established web archiving as a permanent national library service. In the dissertation I outline our initial work plan and how far we managed to get in relation to the formation of the individual objectives, workflows, and professional activities. All the main areas of activities that had to be created in connection with the establishment of the permanent service are being defined.

The third important topic included in the hypotheses is the clarification of the tasks of web librarians and web curators, definition limits of web archiving as a library task in general. In this context, in the one hand, I highlighted the most important library aspects of web archiving in a broader public collection interpretation framework, and on the other hand, the practical workflows discussed in the dissertation. This analysis provides good examples of the appearance of traditional library activities such as acquisition, document processing, document description by format and content, storage, and preservation in web archiving context. I also analyze how the service dimension can be defined in relation to the web archive as an integrated part of the general library collection.

In connection to the fourth topic included in the hypotheses, the evolution of the development of the regulatory and legal background is also being presented as a very important element. Throughout the evolution of the web archiving service, the challenges arising from the lack of infrastructural, legal, and human resources in connection with the establishment of a permanent library service were relevant. Thanks to the open and mostly helpful national library managers and the positive attitude of the relevant ministry officials and legislators, many of the challenges raised in the hypotheses have been overcome. However, in order to carry out the web archiving activity in a sustainable way, it is absolutely necessary to expand the human resources, to maintain the financing background currently guaranteed in the initial period, to continuously ensure the information technology settings, and to cooperate with institutional partners.

The educational issues are being related to both the theoretical background and the foundation of the practical web archiving activities. This subject is being included in the fifth topic of the hypotheses. The topic of web archiving has been successfully included in the framework of accredited librarian training in Hungary in cooperation with the Hungarian Library Institute. An important future task in this field is to build relationships with higher education partners so that theoretical and practical web archiving activities can also be supported in the graduate levels. Followed by the summary of the answers to the hypotheses, I am making a brief evaluation of my publishing activity related to the dissertation. Related publications and conference presentations in the dissertation have been assigned to each main chapter by citations. In connection with the workflows of the web archiving pilot project, I am focusing on several topics such as international web archiving practice, presenting the public collection framework of web archiving, providing metadata for web archives, teaching and using web archiving for scientific purposes, using semantic technologies. By various publications I am presenting the results of the pilot project. I also summarize the content offering of a new website created as a representation of the permanent library service (which also provides an overview of all relevant practical web archiving and related communication activities) and an outline of the software that anyone can use for personal web archiving.

Finally, I would like to mention that the development of the Internet has several open questions, the answer to which could have a fundamental impact on how content on the Web can be archived in the future. A kind of accelerating fragmentation can be observed, on the one hand at the level of different platforms, and on the other hand as a virtual projection of great power rivalry. If the unity of the World Wide Web were to be disappeared even more than before, it could also create a new situation from the point of view of archiving.

The subject of web archiving as a business activity has so far gained very limited space. Should this area become more commercially valuable in the future, and large business players may appear with their services, this may involve a reassessment of the activities of public collections and other non-profit actors that have been prominent on the scene so far.

In this way we have vague ideas about what the future may bring. One thing is for sure, though. The web archiving collection building activity launched in 2017 in Hungary will soon reach a critical level, and the rapidly expanding material of the web archive as a collection will hopefully attract the interest of more and more researchers and offer the opportunity to develop diverse research collaboration opportunities.

Theses of the dissertation

[T1.] A growing segment of our cultural heritage is appearing on the World Wide Web. Collecting, processing, and preserving it in a longer term is a complex challenge. Addressing this task has a few important components in the fields of information science, informatics, social sciences and humanities.

[T2] If web imprints of various fields of everyday life cannot be preserved and presented in public collections, this public institutional system will gradually become incapable of inheriting culture and will also lose its ability to reflect on contemporary social problems. Ultimately, therefore, its legitimacy is called into question. Web archiving therefore needs to be addressed at a scientific level due to its very serious social significance.

[T3] In Hungary, web archiving within an organized framework started with a significant time lag compared to the central European countries. Outlining the

scientific challenges related to web archiving and implementing international research and development results within the practical framework of Hungarian web archiving is an actual task. By setting the foundations for a modern web archiving environment, the negative impact of the consequences of lagging behind can be significantly reduced.

[T4] The description of the broader context of web archiving, the presentation of different research ways based on archived materials, together with the outline of the evolution of modern practical frameworks for web archiving in Hungary, make it possible to perform a complex scientific analysis. As an additional important scientific task, it is vital to define future research directions.

[T5] The practical activities related to web archiving, the theoretical research direction and frameworks of this area are being analyzed from the point of view of public collections. The latest international research results are necessary to be presented, combined with personal Hungarian research experiences, in the form of an individual scientific dissertation in Hungarian. Based on the overview of current international research activities, a new interdisciplinary research domain can be established in Hungarian scientific context.

[T6] The critical precondition for the introduction of web archiving as a topic in Hungarian research context is the creation of a gap-filling scientific work covering the entire field of research, together with a summary and analytical presentation of domestic and international scientific and practical efforts. The analytical approach from the public collection point of view, the presentation of the results of the research and development work belongs to the early stages of institutional web archiving in Hungary, together with an outline of the relevant educational, research and practical exploitation aspects can give impetus to further research.

[T7] As a web-librarian employee of the National Széchényi Library, the author wants to establish the general theoretical research background of web archiving in Hungary, and he also takes an active part in the practical implementation of all the workflows related to the discussed area.

[T8] In addition to a general overview of web archiving, research topics by personal interests are also being formulated from the author's perspective (for example, through the connection between the semantic web and web archiving). It is a constantly evolving dynamic field of science, where it is only possible to

outline a snapshot, a more thorough study of each sub-topic will be a permanent scientific task also in the future.



Registry number: DEENK/62/2021.PL
Subject: PhD Publication List

Candidate: Márton Németh
Doctoral School: Doctoral School of Informatics
MTMT ID: 10049742

List of publications related to the dissertation

Hungarian scientific articles in Hungarian journals (6)

1. **Németh, M.:** Az OSZK Webarchívum új honlapjának felépítése és szolgáltatásai.
Kv. kv. kv. 29 (6), 16-26, 2020. ISSN: 1216-6804.
2. Drótos, L., **Németh, M.:** Egyedi mentésekre szolgáló webarchiváló szoftverek.
Kv. kv. kv. 29 (12), 3-11, 2020. ISSN: 1216-6804.
3. **Németh, M.:** Webarchívum mint a tudományos kutatások tárgya.
Tud. műsz. táj. 67 (12), 757-765, 2020. ISSN: 0041-3917.
4. **Németh, M.:** A digitális könyvtárak, a digitális könyvtáros és a jövő perspektívái.
Kv. kv. kv. 27 (8), 7-17, 2018. ISSN: 1216-6804.
5. Horváth, Á., **Németh, M.:** Kulturális közkincsek feltárása nyílt bibliográfiai adatok révén.
Kvt. figy. 63 (2), 196-210, 2017. ISSN: 0023-3773.
6. **Németh, M.:** Nemzetközi körkép a webarchiválás gyakorlatáról.
Kvt. figy. 63 (4), 575-582, 2017. ISSN: 0023-3773.

Foreign language scientific articles in international journals (3)

7. **Németh, M., Drótos, L.:** A blended learning-based curriculum on Web archiving in the national Széchényi library.
DLP. 35 (2), 97-114, 2019. ISSN: 2059-5816.
DOI: <http://dx.doi.org/10.1108/DLP-03-2019-0012>
8. **Németh, M., Drótos, L.:** Metadata Management and Future Plans to Generate Linked Open Data in the Hungarian Web Archiving Pilot Project.
ITLib. 2, 38-44, 2019. ISSN: 1335-793X.
9. **Németh, M., Simon, A.:** Public collections on the Semantic web in a Hungarian context.
ITLib. 2, 66-74, 2016. ISSN: 1335-793X.





Other journal articles (7)

10. **Németh, M.**: A könyvtári szemantikus webes fejlesztések világa: gondolatok az SWIB konferenciasorozat kapcsán.
Tud. műsz. táj. 66 (3), 157-160, 2019. ISSN: 0041-3917.
11. **Németh, M.**: Webarchiválás két szakmai rendezvény tükrében.
Kv. kv. kv. 28 (26), 29, 2019. ISSN: 1216-6804.
12. Dancs, S., Koltay, T., **Németh, M.**: Adat- és tudásmenedzsment a posztigazság korában: mozaikok az INFINT Information Interactions 2018 nemzetközi konferenciáról.
Kv. kv. kv. 27 (11), 39-43, 2018. ISSN: 1216-6804.
13. **Németh, M.**: A webarchiválás nemzetközi környezete: mozaikok az IIPC 2019 kongresszusról.
Kv. kv. kv. 27 (12), 23-27, 2018. ISSN: 1216-6804.
14. **Németh, M.**: A webarchiválásról történeti megközelítésben: [Ismeretítés].
Kv. kv. kv. 27 (2), 48-52, 2018. ISSN: 1216-6804.
15. **Németh, M.**: Hogyan tudjuk fejleszteni a webes gyűjteményünket?: a Holland Nemzeti Könyvtár webarchiválási tevékenységének értékelése (2007-2017) : [Ismeretítés].
Kvt. figy. 64 (4), 634, 2018. ISSN: 0023-3773.
16. **Németh, M.**: "404 Not Found - Ki őrzi meg az internetet?": webarchiválás workshop az Országos Széchényi Könyvtárban.
Tud. műsz. táj. 64 (11), 577-582, 2017. ISSN: 0041-3917.

Hungarian conference proceedings (3)

17. **Németh, M.**: A mikroadatok felhasználása webarchiválási környezetben.
In: *Workshop 2020 Online : Országos Online Konferencia : 2020. szeptember 2-4.*. Szerk.: Tick József, Kokas Károly, Holl András, Hungamet, Budapest, 81-87, 2020. ISBN: 9786150103761
18. Drótos, L., **Németh, M.**: A webarchiválás oktatása.
In: *Workshop 2018 konferenciakötet.* Szerk.: Tick József, Kokas Károly, Holl András, Hungamet, Budapest, 31-37, 2018.
19. Drótos, L., **Németh, M.**: Az OSZK-ban folyó kísérleti webarchiválási projekt első évének tapasztalatai.
Tud. műsz. táj. 65 (7-8), 389-400, 2018. ISSN: 0041-3917.

Foreign language conference proceedings (5)

20. **Németh, M.**: Using semantic microformats for web archiving: an initial project conception.
In: *CDA 2019 Konferenciakötet, [S.n.]*, Pozsony, 31-38, 2019.





21. **Németh, M.**, Drótos, L.: Web Archives as a Research Subject.
In: Information and technology transforming lives: connection, interaction, innovation : Proceedings of the XXVII. BOBCATSS Symposium. Ed.: Gordana Gašo, Mirna Gilman Ranogajec, Jure Žilić, Madeleine Lundman, University of Osijek, Osijek, 471-478, 2019. ISBN: 9789533141213
22. Drótos, L., **Németh, M.**: The education of web archiving.
In: CDA 2018 Trvalá udržateľnosť a perspektívy ďalšieho rozvoja LTP archívov. Ed.: Tomková Katarína, Univerzitná knižnica v Bratislava, Bratislava, 108-116, 2018.
23. **Németh, M.**, Drótos, L.: Web-Archiving: responsibility of Public Collections in the Preservation of the Digital Culture.
In: Bobcatsss 2018 Riga : The power of reading : proceedings. Eds.: Lelde Petrovska, Baiba Ivane-Kronberga, Zane Meldere, University of Latvia Press, Riga, 124-126, 2018. ISBN: 9789934183539
24. **Németh, M.**: Semantic web developments in Hungarian public collections in an international perspective.
Libellarium. 9 (2), 169-186, 2016. ISSN: 1846-8527.
DOI: <http://dx.doi.org/10.15291/libellarium.v9i2.262>

Foreign language abstracts (2)

25. **Németh, M.**, Drótos, L.: How to catalogue a web archive? Some solutions for metadata management at the web harvesting pilot project of National Széchényi Library, Hungary.
In: Information Interactions 2018 : International conference on digital environment and information science : Book of Abstracts. Ed.: Jela Steinerová, Katarína Buzová, Comenius University in Bratislava, Bratislava, 10, 2018. ISBN: 9788022346047
26. **Németh, M.**, Drótos, L.: Hungarian web archiving pilot project in the National Széchényi Library.
In: 2017 8TH IEEE International Conference on Cognitive Infocommunications : CogInfoCom, Institute of Electrical and Electronics Engineers Inc, Piscataway, 209-211, 2017, (International Conference on Cognitive Infocommunications, ISSN 2375-1312) ISBN: 9781538612644

List of other publications

Foreign language international book chapters (2)

27. **Németh, M.**: Folk culture, literature and nation building efforts in the Romantic Age of Central Europe.
In: Integrity in integration : Developing sustainable dialogue / Rachael Weber, Nagypál Szabolcs, Sajda Peter, Oikumené Akadémická YMCA, Prága, 44-50, 2008, (Ecumenical Anthology ; 7.) ISBN: 978963869469





28. **Németh, M.:** Nationalism, Xenophobia, Populism and Nation States in the Modern History of Central Europe.
In: Roots of Vision, Routes of Change Nationalism and Ecumenical Leadership. Eds.: Rachael Weber, Peter Sajda, Nagypál Szabolcs, BGÖI & WSCF-CESR, Budapest-Bratislava, 34-43, 2007. ISBN: 9789638694652

Hungarian scientific articles in Hungarian journals (17)

29. **Németh, M.:** De hát hol a könyvtár mely célhoz vezet?: az északi országok könyvtáruge a társadalmi mobilitás és kohézió szolgálatában.
Észak. 3 (1), 7-20, 2020. ISSN: 2560-2519.
30. Békésiné Bognár, N. E., **Németh, M.:** Hogyan lépünk egyről a kettőre?: a K2 műhely mltja és jelene.
Kvt. figy. 30 (4), 624-636, 2020. ISSN: 0023-3773.
31. **Németh, M.:** Nemzeti nyelvű könyvtári e-könyvszolgáltatási keretrendszerek: skandináv esettanulmányok itthoni tanulságokkal.
Tud. műsz. táj. 64 (5), 247-251, 2017. ISSN: 0041-3917.
32. **Németh, M.:** Online oktatási tartalomszolgáltatási projekt Nagy Britanniában: a BBC, az oktatási és a kulturális, közgyűjteményi szféra közös platformon.
Kv. kvt. kvt. 26 (2), 30-32, 2017. ISSN: 1216-6804.
33. Bánki, Z., Mészáros, T., **Németh, M.**, Simon, A.: Azonos személyekre vonatkozó név besorolási rekordok automatikus felderítése a PIM adatbázisában.
Tud. műsz. táj. 63 (12), 471-477, 2016. ISSN: 0041-3917.
34. Berger, Á., **Németh, M.**, Simon, A.: Az Europeana Inside és magyar megvalósítása.
Tud. műsz. táj. 61 (9), 342-349, 2014. ISSN: 0041-3917.
35. **Németh, M.:** A Digital Library Learning nemzetközi mesterképzésről - hallgatói szemmel.
Kv. kvt. kvt. 22 (9), 46-52, 2013. ISSN: 1216-6804.
36. **Németh, M.:** Bibzoom - a dán könyvtárak kulturális, zenei szolgáltató-ismeretterjesztő portálja.
Kv. kvt. kvt. 22 (12), 42-45, 2013. ISSN: 1216-6804.
37. Horváth, S. D., **Németh, M.:** Bibliotheca pro futuro.
Kv. kvt. kvt. 20 (10), 21-24, 2011. ISSN: 1216-6804.
38. **Németh, M.:** Közönségkapcsolatok, könyvtári márkanév, civil munka és a szakmai tartalom egyensúlya az észak-európai országokban.
Kv. kvt. kvt. 20 (7), 39-42, 2011. ISSN: 1216-6804.
39. **Németh, M.:** Stratégiai tervezési körkép az észak-európai könyvtárugyről.
Kv. kvt. kvt. 19 (10), 35-42, 2010. ISSN: 1216-6804.
40. **Németh, M.:** Trendek és szerzői jogi problémák a digitális médiaszolgáltatás világában.
Kvt. figy. 56 (2), 263-268, 2010. ISSN: 0023-3773.





41. **Németh, M.**: Áttekintés a modern online audiovizuális szolgáltatások világáról.
Kvt. figy. 55 (1), 29-36, 2009. ISSN: 0023-3773.
42. **Németh, M.**: Az észak-európai országok közönyvtárügyéről.
Kv. kvt. kvt. 18 (1), 48-55, 2009. ISSN: 1216-6804.
43. **Németh, M.**: Az Europeana és a TELPlus projekt.
Kv. kvt. kvt. 18 (3), 3-7, 2009. ISSN: 1216-6804.
44. **Németh, M.**: Az e-könyvek szolgáltatása profitorientált dokumentumticipi körülmények között.
Kvt. figy. 54 (2), 234-240, 2008. ISSN: 0023-3773.
45. Horváth, Á., **Németh, M.**: Az OSZK új közösségi világhálós (web 2.0) fejlesztései.
Kv. kvt. kvt. 17 (8), 34-37, 2008. ISSN: 1216-6804.

Foreign language scientific articles in international journals (2)

46. Bánki, Z., Mészáros, T., **Németh, M.**, Simon, A.: Checking the identity of entities by machine algorithms: the next step to the Hungarian National Namespace.
Code4Lib. 33, 11765, 2016. ISSN: 1940-5758.
47. **Németh, M.**: Hungarian Libraries and Librarianship, 1990-2013: an Overview.
Library Trends. 63 (2), 212-232, 2014. EISSN: 1559-0682.
DOI: <http://dx.doi.org/10.1353/lib.2014.0034>

Other journal articles (27)

48. **Németh, M.**: Információs műveltség és a könyvtárak a hamis hírek korszakában.
Kvt. figy. 66 (3), 507-510, 2020. ISSN: 0023-3773.
49. **Németh, M.**: Szubjektív mozaikok a Networkshop 2020 konferenciáról.
Kvt. figy. 30 (4), 668-675, 2020. ISSN: 0023-3773.
50. **Németh, M.**: 85. IFLA World Library Information Congress 2019, Athén - rövid szubjektív áttekintés.
Könyvtárvilág. 7 (4), 1, 2019. EISSN: 2064-1443.
51. Ford. **Németh, M.**: Áttekintés a dán könyvtárosok adattudományt érintő kompetenciáiról és gyakorlati tevékenységeiről.
Kvt. figy. 65 (4), 611-615, 2019. ISSN: 0023-3773.
52. Arany, Z., Nagy, A., **Németh, M.**, Radó, R.: Mozaikok az eszéki Bobcatssz 2019 konferenciáról.
Kv. kvt. kvt. 28 (2), 16-23, 2019. ISSN: 1216-6804.
53. **Németh, M.**: Színes körkép egy összetett könyvtári területről.
Kvt. figy. 65 (4), 616-618, 2019. ISSN: 0023-3773.
54. **Németh, M.**: A REACH projekt: közösségi platform a társadalmi részvétel alapuló tevékenységek, illetve a közösségi innováció elősegítésének céljából.
Kv. kvt. kvt. 27 (5), 18-20, 2018. ISSN: 1216-6804.





55. **Németh, M.**: Digitális Bölcsészet. Gondolatok egy új nyílt hozzáférést, interdiszciplináris e-folyóirat születése kapcsán.
Kv. kv. kv. 27 (12), 48-52, 2018. ISSN: 1216-6804.
56. **Németh, M.**, Nagy, A.: Mozaikok a rigai Bobcatss 2018 konferenciáról.
Kv. kv. kv. 27 (2), 11-21, 2018. ISSN: 1216-6804.
57. **Németh, M.**: Szubjektív beszámoló az IFLA wroclawi világkonferenciájáról.
Kvt. figy. 64 (2), 273-278, 2018. ISSN: 0023-3773.
58. Arany, Z., Bálint, D., Farkas, C., Kállai, A., Nagy, A., **Németh, M.**, Tóth, V.: A 25. jubileumi BOBCATSSS konferencia az ezer tó országában.
Kv. kv. kv. 26 (2), 38-41, 2017. ISSN: 1216-6804.
59. **Németh, M.**: Áttekintés a Szakkönyvtári Seregszemle 2017 konferenciáról.
Kv. kv. kv. 26 (6), 21-26, 2017. ISSN: 1216-6804.
60. Ford. **Németh, M.**: Beszélgetésre épülő programok svéd és norvég közönyvtárakban: esettanulmányok a közönyvtárak integrációban betöltött szerepéről.
Kvt. figy. 63 (Különszám), 83-89, 2017. ISSN: 0023-3773.
61. Virág, G., Lendvay, M., **Németh, M.**: Mozaikok az ILIDE 2017 konferenciáról.
Kv. kv. kv. 26 (7), 23-28, 2017. ISSN: 1216-6804.
62. Csikász-Nagy, Á., **Németh, M.**, Pataki, F., Tót, B.: Önkéntesként szerzett benyomások 2017 kiemelkedő nemzetközi szakmai eseményéről: IFLA World Library and Information Congress 83rd IFLA General Conference and Assembly 2017. augusztus 19-25. Wroclaw, Lengyelország.
Kv. kv. kv. 26 (11), 11-17, 2017. ISSN: 1216-6804.
63. Arany, Z., Jávorka, B., Nagy, A., **Németh, M.**: Egy hét az oroszlan barlangjában: BOBCATSSS 2016, Lyon.
Kv. kv. kv. 25 (3), 17-22, 2016. ISSN: 1216-6804.
64. **Németh, M.**: Mozaikok a göteborgi könyvvásárról.
Kvt. figy. 62 (2), 245-250, 2016. ISSN: 0023-3773.
65. **Németh, M.**: Mozaikok az információs műveltség európai konferenciájáról (ECIL, 2015).
Kvt. figy. 62 (1), 86-90, 2016. ISSN: 0023-3773.
66. **Németh, M.**: Oslo, Helsinki, Lyon, Brno: konferencia-mozaikok 2014 tavasza és 2015 januárja között.
Kvt. figy. 61 (1), 75-84, 2015. ISSN: 0023-3773.
67. **Németh, M.**: Szemelvények a MoodleMoot 2015 konferenciáról.
Kv. kv. kv. 24 (9), 28-33, 2015. ISSN: 1216-6804.
68. **Németh, M.**: E-könyvekről sokszemközt: Az ELTE BTK Könyvtár- és Információtudományi Intézetének szakmai napja.
Kvt. figy. 60 (4), 486-491, 2014. ISSN: 0023-3773.



