

VALÓS TÉRBEN – AZ ONLINE TÉRÉRT

Networkshop 31: országos konferencia

2022. április 20–22.
Debreceni Egyetem

Szerkesztette: Tick József, Kokas Károly, Holl András

HUNGARNET Egyesület
Budapest, 2022



A kötet megjelenését támogatta az
Energiaügyi Minisztérium

Szerkesztette: Tick József, Kokas Károly, Holl András

Tipográfia és tördelés: Vas Viktória

Workshop

2022. április 20–22. Debreceni Egyetem, konferencia előadásainak közleményei

ISBN 978-615-82243-0-7

DOI: [10.31915/NWS.2022](https://doi.org/10.31915/NWS.2022)

Kiadja a HUNGARNET Egyesület
az MTA Könyvtár és Információs Központ közreműködésével
Budapest
2022

Borítókép: [freepik.com](https://www.freepik.com)

TARTALOMJEGYZÉK

Előszó	5
Lencsés Ákos: A nyílt tudomány pénzügyi vonatkozásai	7
Farkas Katalin: Centenáriumi média-adattár és virtuális kiállítás létrehozásának tanulságai az SZTE Klebelsberg Könyvtárban	13
Bódog András: A nyílt archívumi információs rendszer (OAIS) szabványának honosítása.....	20
Perlaki Attila: Oktatást segítő gamifikációs alkalmazások, mint szakdolgozati témák	27
Csapó Noémi – Dani Erzsébet: APPropó fejlődés – A Bács-Kiskun Megyei Katona József Könyvtár mobilapplikációja.....	32
Simon András: Integrált könyvtári rendszerek tranzakciós rekordjainak vizsgálata, a könyvtári állomány digitalizálásának tervezésekor.....	41
Németh Márton: Az OSZK Webarchívum nemzetközi kapcsolatai.....	58
Antal Péter: A mesterséges intelligencia kihívásai a XXI. század társadalmára	70
Hajdu Csaba – Szilágyi Zoltán: Modern robotikai technológiai ismeretek oktatása „Teljes spektrumú” oktatási módszerrel	77
T. Nagy László – Boda István Károly – Tóth Erzsébet: E-tananyagfejlesztés virtuális 3D környezetben.....	84
Palencsárné Kasza Marianna: Digitális átállás – Minőség – lehetőségek az EQAVET terén.....	92
Nagy Gyula: Nemzetközi kitekintés a felsőoktatási könyvtárak világára: a EUGLOH könyvtári workshopja	99
Babocsay Gergely: Az európai természettudományi gyűjtemények digitális integrációja: határ a csillagos ég.....	108
Somorjai Noémi: Egyenlőtlenségek a tudományos kutatás területén. Az amatőr kutatók szerepe	114
Molnár Dániel – Dani Erzsébet: Robotok a könyvtárban: Hogyan válhat a robotika a könyvtári mindennapok részévé?	122
Horváthné Felföldi Helga: Digitalizáció a szakképzésben. A Szakmajegyzékben szereplő szakmák digitáliskompetencia jártassági szintjeinek felülvizsgálata	130
Kalcsó Gyula: Ne csak útra csomagoljunk! Miért fontos a csomagolás a digitális megőrzésben?	138
Karsa Zoltán István – Szeberényi Imre: A CIRCLE felhő elmúlt évtizede	146
Bobák Barbara – Kasza Péter: Az MI lehetőségei a kora újkori filológiában: Johannes Michael Brutus <i>Rerum Ungaricarum</i> libri kéziratának digitális kiadása (esettanulmány)	154
Egyed-Gergely Júlia – Vajda Róza, Gárdos Judit – Horváth Anna – Meiszterics Enikő – Micsik András – Martin Dániel – Marx Attila – Pataki Balázs – Siket Melinda: Szociológia, kutatási adatok, mesterséges intelligencia: lehetőségek és tapasztalatok	161
Szemes Botond – Bajzát Tímea – Fellegi Zsófia – Kundráth Péter – Horváth Péter – Indig Balázs – Dióssy Anna – Hegedüs Fanni – Pantyelejev Natali – Sziráki Sarolta – Vida Bence – Kalmár Balázs – Palkó Gábor: Az ELTE Drámakorpuszának létrehozása és lehetőségei.....	170



Sebestyén Ádám: Az ELTEdata szemantikus adatbázis legújabb fejlesztései.....	179
Szlamka Erzsébet: Új trendek a tanulási eredmények tanúsításában	185
Tóth Máté – Héjja Balázs: Webshop indítása közkönyvtári környezetben.....	192
Etlinger Mihály – Hernády Judit: A kiadás hagyatéka / a hagyatéka kiadása: A Régi Magyar Költők Tárának hálózati kiadásáról.....	199
Varga Emese – Makkai T. Csilla: „Ki a fenének kell collstok?” A digitális szöveg rejtett mértékegységei	204
Dobás Kata – Fazekas Júlia: ITIdata – Egy irodalmi adatbázis fejlesztése Wikibase alapon és ennek hasznosítása Kosztolányi Dezső forrásjegyzékénél	211
Sörény Edina: Kézai Simon Program – digitális családi fotóarchívum.....	219
Fülöp Tiffany – Molnár Tamás – Hoczopán Szabolcs: Open Monograph Press e-könyvplatform a Szegedi Tudományegyetemen	227
Palkó Gábor: Mesterséges intelligencia, digitális bölcsészet, kulturális örökség: trendek és eredmények.....	235
Pergéné Szabó Enikő – Bátfai Mária Erika: A tudományos publikálás támogatása a Debreceni Egyetemi és Nemzeti Könyvtárban	241
Csirmazné Rezi Éva: Nemzetközi kiadványazonosítók és kötelezpéldányok kezelése az OSZK OKP (Országos Könyvtári Platform) rendszerében	250
Alföldi István – Dióssy Anna Laura: Digitálisan született kutatási anyagok megőrzése: a relációs adatbázis mint born-digital objektum	262
Fekete Norbert: HTR-modellépítés és kézírásfelismerés nagyméretű, többszerzős szövegtörzsen. A Transkribus alkalmazása az Arany János hivatali iratokon.....	271
Horváth Péter – Kundráth Péter – Palkó Gábor: ELTE Népdalkorpusz – magyar népdalok gépileg annotált adatbázisa	276
Nagy György: IKT eszközök alkalmazása az alsó tagozatos környezetismeret órákon.....	284
Köpösdí Zsuzsa – Molnár Tamás: Multimédiás, interaktív és adaptív tananyagok létrehozásának lehetőségei H5P keretrendszerrel	289
Jankó Tamás: Munka 4.0 – Ipar 4.0 – Szakképzés 4.0 – : A digitális kompetencia jövőbeni fejlesztési útjai	296
Békésiné Bognár Noémi Erika – Nagy Andor: Megújuló könyvtári statisztika: az egységes adatstruktúra és a korszerű megjelenítés kialakításának útján	304
Bolya Máttyás: Kézírtos dallamlejegyzések feldolgozása MI-vel támogatott digitális környezetben	310
Maróthy Szilvia – Seláf Levente – Vigyikán Villó: Régi magyar verskorpusz összeállítása stilometriai és számítógépes metrikai kutatásokhoz	324
Szűcs Kata Ágnes: Kézírtos források transzformációinak lehetőségei a közgyűjteményekben.....	330
Fellegi Zsófia: A digitális filológia infrastruktúrái. A DigiPhil megújulásáról.	338
Mihály Eszter: Mi az a dHUpla? A Digitális Bölcsészeti Platform bemutatása.....	345
Nemeskey Dávid Márk – Palkó Gábor: Szemantikus névelém-azonosítás magyar nyelvű szövegeken (a HuWikifier bemutatása)	359

Az OSZK Webarchívum nemzetközi kapcsolatai

Németh Márton
Országos Széchényi Könyvtár
Digitális Bölcsészeti Központ
Digitális Filológiai és Webarchiválási Csoport
ORCID: [0000-0003-1864-8107](https://orcid.org/0000-0003-1864-8107)

Az alábbiakban rövid összefoglalást nyújtunk az OSZK webarchívumának külföldi kapcsolatairól, illetve a nemzetközi munkánk során elért eredményekről. A tanulmányban öt különféle témakörrel ejtek szót. Elsőként a webarchiválási projektünk megalapozásában kulcsszerepet játszó, dániai szervezésű online tanfolyamról ejtek szót. Ezt követően bemutatom az OSZK részvételét a webarchiválásért felelős nemzetközi konzorcium (IIPC) szakmai tevékenységeiben, különös tekintettel az oktatási munkacsoportra. A harmadik téma a közép-európai webarchívum együttműködési kezdeményezés (CEWA), melyet a 2021-ben megrendezett „404 Not Found” online workshop első napjának délutánján tartott kerekasztal-beszélgetésén alapoztuk meg. Felvázolom az együttműködés tervezett területeit, s a várható eredményeket. A webarchívumok kutatási célú felhasználásának kérdéseire fókuszál a 2023 nyaráig tartó WARCNET projekt, mely a dániai Aarhusi Egyetem koordinálásában valósul meg. Negyedik témaként röviden összefoglalom a projekt főbb területeit, eddigi eredményeit, s a jelenleg is zajló tevékenységekről is áttekintést nyújtok. A zárótémakör keretei között pedig az amerikai Internet Archive magyar állományára alapozódó teljesszövegű kereső és portál fejlesztéséről adok áttekintést.

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, 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 Consortium-hoz, aktívan bekapcsolódtunk több munkacsoport tevékenységébe is. Emellett felvettük a kapcsolatot közép-európai partnerintézményekkel későbbi intenzívebb együttműködési formák megalapozása céljából.

1. A NetLab kurzus

A dániai Aarhus egyetemén az országos DIGHUMLAB kutatási infrastruktúra részeként működő, Niels Brügger professzor vezette NetLab kutatócsoport online egyetemi kurzusokat kínál PhD hallgatók, ill. a témában érintett szakemberek, kutatók számára a webarchiválás témakörében. 2017 őszén a webarchiválási projektünk indulásakor első alkalommal – angol nyelven – nemzetközi keretek között is meghirdették a kurzust, ráadásul a részvétel ingyenes volt. A szervezési-moderátori feladatokat Asger Harlaug, a kutatócsoport helyettes vezetője látta el. A szakmai tartalom mellett nagyon fontos volt a részvétel abból a szempontból is, hogy áttekintést kaptunk arról, hogy miként kell egy webarchiválással foglalkozó tanfolyamot felépíteni, megszervezni és gondozni. Szembesülhettünk a távoktatási felületen zajló kollaboratív munka módszertani jellegzetességeivel is. Külön növelte a kurzus értékét, hogy az OSZK kísérleti projektjével nagyjából egyidőben indult belga webarchiválási program munkatársaival is alkalmunk nyílt megismerkedni, a dán szakemberekkel zajló közös munka mellett. Mély benyomást tett a résztvevőkre a konstruktivista pedagógiai háttér, a segédanyagok, a kurzus menetrendje, a közös munkafelület. Emellett azonban alapvetően a résztvevők szakmai irányultságán, lelkesedésén múlt, hogy milyen tartalommal lehetett

meztölteni a közös munkát. Végső soron a tanfolyam sikerének talán kisebbik hányadát adta az előzetes felkészülés, a tananyag kidolgozása. A sikeresség a tartalom megfelelő felvázolásán, a résztvevőkkel zajló osztálytermi, illetve virtuális kommunikáció minőségén állt vagy bukott.

Az online szeminárium egyéni felhasználó név/jelszó párossal védett, Moodle-alapú távoktatási felületen zajlott. Itt publikálták a szervezők a kurzus segédanyagait, a feladatok szövegeit, és itt volt lehetőség a válaszok megfogalmazására és az egymással történő társalgásra is a kurzus-fórum keretében. Ez a felület csak a tanfolyam ideje alatt volt elérhető, de annak végeztével lehetőségünk volt az összegyűlt anyag lementésére.

Öt fő terület került megtárgyalásra. Először annak felmérése zajlott, hogy a résztvevőknek milyen kapcsolatuk van a webarchiválással és milyen elvárásaik fogalmazódnak meg a kurzussal szemben? Már ebből is következtetni lehetett arra, hogy a konstruktivista pedagógiai alapon szerveződő szeminárium minden egyes félévben más és más profillal bír, a tartalmát erőteljesen befolyásolja a résztvevők érdeklődési köre, szakmai irányultsága.

A második téma körül zajló társalgás éppen arra irányult, hogy az általános érdeklődési kör megfogalmazása után mindenki elmondja, hogy miért hasznos számára webarchiválással foglalkozni és felvázoljon valamiféle, a kurzus további teljesítéséhez kapcsolódó konkrét munkatervet vagy kutatási tervet.

A harmadik témakörben keresni kellett három, az egyéni érdeklődési területhez tartozó honlapot, melyek legalább egy éve üzemelnek már, s áttekintést kellett adni ezek archiválásának kihívásairól. Mintaként lehetett használni az Internet Archive vagy bármelyik nemzeti webarchívum nyilvánosan elérhető mentéseit. Itt lehetett igazán először megtapasztalni a webarchiválásra jellemző technikai-szakmai kihívások sokaságát és szinte áttekinthetetlen összetettségét. A kurzus résztvevői a saját munkájuk során szembesülhettek azzal, hogy egyáltalán milyen keretek között lehet webarchiválásról beszélni, hogy az archiválás sikerességének kritikus tényezője a megfelelő szoftveres eszköztár kiválasztása, mely sokszor csupán a tartalom szerkezetének alapos feltérképezése után végezhető el. Persze szembenézhattünk a korlátozott szakmai-emberi erőforrások természetével is, hogy a webarchiválási tevékenység szempontjából kritikus jelentőségű a rendelkezésre álló kapacitások hatékony felhasználásának művészete is.

A negyedik témakörben a kutatási terület (gyűjtőkör) előzetes felmérése, valamint a megadott idő alatt egy adott személy által elvégezhető munka kereteinek operatív meghatározása után meg kellett fogalmazni a kutatási célunkhoz kötődő további honlapok összeválogatásának stratégiáját. Ezt követően ki kellett választani a megfelelő szoftveres háttérrel, kísérleti aratásokat végezni, majd mini riportban összegezni a tapasztalatokat (Németh, 2018c). Számomra ez bizonyult a leghasznosabb pontnak, hiszen a belgák is és mi is elő tudtunk állni konkrét tartalmak kísérleti mentéseinek tapasztalataival, s meg tudtuk beszélni egymással az aratáshoz használt szoftverek előnyeit, hátrányait, illetve értékeltük a végeredményt. A társalgásba több ponton bekapcsolódtak a dán nemzeti webarchívum munkatársai is, akik saját tapasztalataikkal színesítették a kommunikációt. Az volt a fő cél, hogy mindenki választ kapjon a kérdéseire a konkrét kihívások kapcsán, s ezeket hasznosítva tudja tovább építeni a saját projektjét.

Mindezek után pedig az ötödik témakör keretében összesítettük a kurzussal kapcsolatos tapasztalatainkat, s ezzel zárult a szeminárium, melynek elvégzéséről a résztvevők bizonyítványt is kaptak.



2. Az IIPC konzorcium és az abban zajló magyar tevékenység

Az egyes nemzeti webarchiválási modellek (könyvtárakban, levéltárakban, egyetemeken, kutatóintézetekben) tulajdonképpen arra válaszul jöttek létre, hogy a web fejlődése egyetlen szervezet számára már követhetetlené vált archiválási szemszögből. Ezek egy része eleve csak korlátozott időre szólt, másokat pedig egy idő után újragondoltak és -terveztek, így már több második-generációs webarchívummal is lehet találkozni. A nemzeti szereplők, illetve az Internet Archive egymást kiegészítő tevékenységei adják meg a webarchiválási szolgáltatások teljességét. Jelenleg mintegy 45 nemzeti webarchívum létezik valamivel több mint harminc országban, ugyanis egyes helyeken az államalkotó nemzeti népcsoportoknak külön archívumuk van. Ezeket a webarchiválási tevékenységeket az IIPC nemzetközi konzorcium fogja közös keretbe, melynek az Országos Széchényi Könyvtár is tagja lett 2018-ban. Ennek révén hatékony tudásmegosztási, kompetenciafejlesztési és szakmai együttműködési csatornák válnak elérhetővé számunkra is. A konzorcium 2003-ban alakult meg a Francia Nemzeti Könyvtár kezdeményezésére 12 alapító taggal. Jelenleg a 45 országból összeálló tagság könyvtárakat, levéltárakat, múzeumokat, illetve a webarchiválás területén érdekelt civil és piaci szereplőket is magában foglal. A konzorcium alapvető küldetése az internetes tartalmak archiválásához és hosszú távú biztonságos megőrzéséhez szükséges tevékenységek koordinálása és támogatása¹.



1. ábra: Az IIPC honlapja

Az IIPC tevékenységét korábban a British Library-ből koordinálták, manapság már önálló titkárságot működtetnek. Az operatív irányítást a tagintézmények képviselőiből választott Végrehajtó Bizottság látja el. A közös munka nagy része tematikus munkacsoportokban

1 <https://netpreserve.org>

zajlik, melyek jelenleg a tartalomfejlesztéssel, a hosszú távú megőrzéssel, illetve az oktatási kérdésekkel foglalkoznak. Az OSZK képviselőjében ez utóbbi (2017 végén alakult) munkacsoportban, illetve a 2019-ben indult kutatási munkacsoportban veszünk részt. Külön figyelmet érdemel a 2019-2020-ban létrehozott repozitórium², melynek keretében fokozatosan teszik közzé az IIPC szakmai rendezvényein elhangzott előadások prezentációit, a kutatási jelentéseket és cikkeket. Ennek használata szinte azonnal megkerülhetlenné vált a webarchiválás tanulmányozásához. Az aktuális kutatási területek jóval rövidebb átfutási idővel bukkannak itt fel, mielőtt tudományos publikációkban jelennének meg. Képet kaphatunk arról, hogy mely témák állnak a gyűjteményfejlesztéssel, a technikai fejlesztésekkel, illetve a webarchívumok kutatási célú felhasználásával foglalkozó szakemberek érdeklődésének középpontjában.

Alkalmam nyílt részt venni az IFLA 2017. évi kongresszusán a lengyelországi Wrocławban, ahol önálló szekció foglalkozott az IIPC szervezésében a webarchiválással. Itt történt meg 2017 augusztusában az első személyes kapcsolatfelvétel az IIPC tagjaival, vezetésével. Tanácsokat kaptam tőlük a webarchiválási kísérleti projektünk megalapozásához, még mielőtt hivatalosan taggá váltunk volna. Az IIPC tagságunk lényeges eredménye az a kutatás-fejlesztési együttműködés, melyet szerződéses munkatársunk, Vitéz Gábor alakított ki a Dán Nemzeti Könyvtár munkatársaival, elsősorban Thomas Egense-vel, aki a SolrWayback nevű teljes szövegű kereső és megjelenítő eszközt fejleszti. A szoftver tesztelésébe és egyes hibák javításába mi is be tudtunk segíteni. A dán fejlesztők az IIPC 2018-as, mi pedig az OSZK képviselőjében az IIPC 2019-es közgyűlésén számoltunk be e közös munka eredményeiről, mely a konzorcium vezetésének körében is pozitív visszhangot keltett.

Csatlakozásunk évében még nem volt módunk a IIPC konzorcium éves közgyűlésén, illetve az ahhoz kapcsolódó konferencián való részvételre, mivel ezekre Új-Zélandon került sor. Lehetőségünk nyílt viszont egy öt perces videó összefoglalót készíteni a rendezvényen részt vevők számára, bemutatva magunkat és eddigi eredményeinket. 2019-ben már részt vállaltunk a Zágrábban rendezett közgyűlés szervezésében, a programbizottság munkájában is. A kísérleti projektről tartott rövid összefoglaló mellett a dán partnereinkkel közös fejlesztési munkákról adtam elő, a SolrWayback és a saját fejlesztésű SolrMIA keresőkről, és azok találati listáinak relevánsabbá tételéről metaadatok segítségével. Sikeres együttműködést alapoztunk meg a Web Curator Tool webarchiválási munkafolyamatokat segítő keretrendszer új-zélandi és holland fejlesztőivel, a rendezvényen tartott workshop alatt. 2019 őszén ennek tapasztalatai alapján munkatársaimmal egyeztetve eljuttattam kérdéseinket a programcsomag fejlesztése kapcsán a partnereinkhez, így közvetlen befolyással bírnak, informatikus kollégáinkkal együtt, a fejlesztési irányok meghatározására. Megteremtettem az alapjait a közép-európai webarchívumok közötti együttműködésnek is, melynek konkrét szervezeti kereteit közös fejlesztési programokkal, rendezvényekkel a 2021. év során terveztük partnereinkkel együtt meghatározni.

2 <https://digital.library.unt.edu/explore/partners/IIPC/>

SOLRWAYBACK

The screenshot shows the SolrWayback search interface. The search term 'Múzeumkert' is entered in the search bar. Below the search bar, there are options for 'GROUPED SEARCH', 'IMAGE SEARCH', and 'URL SEARCH'. The 'APPLIED FACETS' section shows the domain 'mnm.hu' selected. The 'RESULTS' section displays a list of facets on the left and a detailed view of a search result on the right. The detailed view shows the title 'MÚZEUMKERT | MAGYAR NEMZETI MÚZEUM', the type 'html, web page @ mnm.hu', the date '20/05-2020', and the URL 'https://mnm.hu/hu/muzeumtortenet/muzeumkert'. The highlighted content is 'Akadálymentes Facebook Instagram Free WiFi Bejelentkezés Látogatók Látogatók Megközelítés Nyitva tartás A Múzeumkert házirendje'. Below the text, there are four image thumbnails and a 'View data fields' button.

2. ábra: A Solrwayback kereső találati oldala

Az IIPC konzorciumon belül 2017 végén alakult meg az oktatási és képzési munkacsoport³. Első projektjük keretében kérdőíves vizsgálatot végeztek. Arra irányult a kutatás fókusz, hogy az egyes országokban milyen típusú szervezetek, mekkora létszámban foglalkoznak webarchiválással, valamint ezeknek a szakembereknek milyen igényeik vannak az oktatás, a szakmai továbbképzés terén. Az IIPC tréning munkacsoportja a kérdőív tapasztalatait felhasználva először felmérte azokat a fajta oktatási tevékenységeket, melyekkel a leghatékonyabb segítséget tudják majd nyújtani tagintézményeik, illetve a lehető legszélesebb körű szakmai célcsoportok számára. Fokozottan szeretnének építeni az egyes országokban eddig felhalmozott tapasztalatokra. Most először mérték fel azt is, hogy az egyes országokban milyen oktatási és képzési tevékenységek zajlanak, milyen keretek között, s hányan vesznek azokban részt. A munkacsoport online értekezletein lezajlott eszmecsere során kiderült, hogy a legtöbb oktatási tevékenység igen szűk körű, csak az egyes archiválási tevékenységek ellátásához feltétlenül szükséges készségek és képességek átadására fókuszáló oktatási tevékenység nyert teret. Olyan széleskörű, általános tananyagok kidolgozása⁴ lett tehát az elsődleges cél, melyek rugalmasan beilleszthetők bármiféle szervezet fizikai jelenlétre épülő, illetve távoktatási programjának a keretei közé, megbízható szakmai tartalmat kínálva fel a webarchiválással kapcsolatos különböző kompetenciák megalapozásához. A későbbiekben kerülhet sor haladóbb szintű webkuratori, illetve informatikai ismeretekhez kötődő témakörökhöz illeszkedő tananyagfejlesztésre is.

3 <http://netpreserve.org/about-us/working-groups/training-working-group/>

4 <http://netpreserve.org/web-archiving/training-materials/>

Aszakmai tartalmak meghatározása mellett a módszertani felépítés is egyenrangú fontossággal bír a tananyagok fejlesztése során. A moduláris konzisztencia megteremtése különféle szakmai csoportokban, szinteken válik ellenőrizhetővé. A csoport általános álláspontjaként lett elfogadva, hogy habár fontos a fejlesztésre fordított időtényező optimalizálása is, de a megfelelő minőségű oktatási anyag rendelkezésre bocsátása élvez elsődlegességet. A tananyagok mögött álló koncepció megvilágítására a zágrábi IIPC konferencia alatt videó interjúk is készültek a webarchiválás oktatásának különféle aspektusairól. A munkacsoportban összeállítottunk egy mintegy száz tételből álló kérdéslisztát. Az interjúalanyok számára a felvétel készítői (előzetes konzultációt követően) ezekből szemezgettek különféle témakörökhöz illeszkedve. Az interjúkat a tananyagok közrebocsátásával egyidőben publikálták⁵.

Az elkészült oktatási anyag az IIPC különféle fórumain keresztül kerül terjesztésre. Egy önálló GitHub csatorna létrehozását⁶, az alapvetően informatikusok által szoftverek közzétételére szolgáló felület használatát több tényező is indokolja. A verziókövetéssel jól nyomon követhetők az oktatási anyagban bekövetkező változások. A kiegészítő vizuális és szoftveres elemek is könnyedén beilleszthetők a tananyag mellé. A felhasználói visszajelzések, korrekciós javaslatok precízen rögzíthetők és dokumentálhatók.

Az első nemzetközi kurzus a Digital Preservation Coalition és az IIPC Training Working Group szervezésében 2020 októberében került megszervezésre, a webarchiválásban még nem jártas IIPC tagintézmények munkatársai részére. Itt kerültek először kipróbálásra a kifejlesztett tananyagok konkrét oktatási tevékenység közben is. A tanfolyam a korábban említett modulok anyagát kínálja, válaszokat keresve arra, hogy miért is kell a webet archiválni, hogyan néz ki egy webarchívum, milyen koncepciók szerint, s milyen technológiai eszközökkel lehet webarchiválást szervezni és lebonyolítani, és hogy milyen kutatói és egyéb felhasználói igények kielégítésére kell megoldásokat találni.

3. Közép-európai együttműködés

A webarchiválási tevékenység elindulásától kezdve személyes találkozásokon alapuló kapcsolatokat sikerült kiépíteni a cseh, a szlovák és az osztrák webarchívumok munkatársaival is. A pozsonyi Egyetemi Könyvtárban több előadást is tartottunk, illetve információcserét folytattunk a házigazda intézmény és a Cseh Nemzeti Könyvtár webarchiválásért felelős munkatársaival. Az Osztrák Nemzeti Könyvtárban pedig egy napos szakmai tanulmányúton vettünk részt munkatársaimmal együtt, hogy megismerjük az ott zajló webarchiválási tevékenységet. Sikeres szakmai rendezvényt is szerveztek fennállásuk 10. évfordulóján, ezen is alkalmam nyílt részt venni, ahol elsősorban a német nyelvű országok webarchiválási projektjeiről lehetett széleskörű áttekintést szerezni. A horvát kollégákkal a 2019-ben megrendezett zágrábi IIPC konferencián kerültünk közelebbi kapcsolatba.

Az OSZK által ötödik alkalommal megszervezett „404 Not Found” online konferencia és workshop első napján (2021. november 23.) bemutakoztak a közép-európai országok nemzeti szintű webarchívumai Ausztriából, Csehországból, Horvátországból, Szlovéniából és Szlovákiából, valamint a Lengyel Állami Levéltár is ismertette a terveit (ott 2022 elején indult el a webarchiválási tevékenység előkészítése). Az angol nyelvű prezentációk elérhetőek a rendezvény weboldalan⁷.

A délutáni kerekasztal beszélgetésen ismertettük, majd megvitattuk a résztvevőkkel azt a javaslatunkat, hogy ezek a webarchívumok működjenek szorosabban együtt, egy Central

5 <http://netpreserve.org/web-archiving/iipc-training-video-case-studies/>

6 <https://github.com/iipc/training>

7 <https://webarchivum.oszk.hu/404-not-found-workshop-2021-november-23-24/>



European Web Archives (CEWA) nevű projekt keretei között, tekintettel a régió közös kulturális és történelmi örökségére, ami az interneten is leképeződik. A kezdeményezést mindenki pozitívan fogadta, de természetesen ez még nem jelenti az anyaintézmények hivatalos csatlakozását a CEWA projekthez, jelenleg az előkészítés zajlik. Terveink között szerepel egy többnyelvű portál és metaadatkereső megalkotása. A portálon a résztvevő nemzetek nyelvén bemutatkozna mindegyik webarchívum. Erre a célra felhasználnánk az IIPC által támogatott „Dark and Stormy Archives” nevű rendszert⁸, mellyel látványos és a közösségi médiában is megosztható módon hozhatók létre válogatások archivált webtartalmakból. Beépítenénk továbbá a szintén IIPC-s forrásból fejlesztett technológiát, a „Bloom filter” módszer használatát a webarchívumokban való közös kereséshez⁹. Ennek az a lényege, hogy a korábban erre a célra használt Memento protokollnál nagyságrendekkel gyorsabban lehet megállapítani, hogy egy, a felhasználó által keresett URL cím megvan-e valamelyik webarchívumban. (A hozzáférés utána természetesen attól függ, hogy egy nyilvános vagy egy csak helyben használható gyűjteményrészben van-e az adott URL-ről lementett fájl.) Ezt a megoldást a horvát webarchívumban tesztelték tavaly, így az ottani kollégáknak már van ezzel tapasztalatuk. Fontos lenne az egymás közötti tapasztalatcsere és a jó gyakorlatok átvétele a webarchiválással kapcsolatos tevékenységek terén (különös tekintettel az informatikai fejlesztésekre). Közös rendezvények és együttes publikációs tevékenységek szervezése is a tervek között szerepel. Az együttműködő intézmények megoszthatnák egymással saját URL listáikat a regionális jelentőséggel bíró témák, események kapcsán és felmerülhet az archiválási feladatok részleges összehangolása is. Emellett közös projektekkel lehetne pályázni az IIPC forrásaira is.

4. A WARCNET projekt

A projekt eredeti célkitűzése egy olyan kutatási hálózat megalkotásában rejlett, mely segít értelmezni a nemzeti, illetve nemzetközi felsőszintű web doménekből archivált anyag felépítését, rétegzettségét történetét, tartalmi összefüggéseit (ilyenek például a .hu vagy a .com tartományok), illetve a globális hatású társadalmi események európai szintű webes lenyomatát igyekszik feltárni, mint a webarchívumokban tárolt digitális kulturális örökség fontos szeletét. Számos társadalmi krízis, konfliktus, vitatéma tanulmányozásához elengedhetetlenek a webes források. A projekt közös platformot nyújt a webarchívumokban tárolt anyagot kutatóknak, illetve a gyűjtemények működtetésében főszerepet játszó közgyűjteményi szakembereknek és informatikusoknak is. A csapat tagjai egyéni kutatók, a nemzeti könyvtári webarchívumok és az IIPC munkatársai is. A fő cél, hogy az archívumokban tárolt anyagokban rejlő összefüggésekre, a web történetére vonatkozó kutatásokra sok szempontból lehessen tekinteni (pl. kommunikáció- és médiatudományi, történettudományi, digitális bölcsészeti nézőpontokból).¹⁰ Az 1990-es évektől a számítógépes világháló fokozatosan a társadalmi kommunikáció, a tudásátadás, illetve tágabb értelemben a közélet megkerülhetetlen médiumává lett, természetéből adódóan nemzetközi térben. Az itt zajló diskurzusok szinte egy pillanat alatt a múlt részévé válnak, ezek feltárása, elemzése nagyon fontos kutatási tevékenységként jelenhet meg. Az Internet Archive, illetve a nemzeti könyvtárak és levéltárak javarészt már évtizedek óta gyűjtik a webes múlt fontos emlékeit. A webarchívumok kutatása azonban számos esetben speciális kompetenciákat igényel, messze nem olyan triviális, mint a hagyományos levéltári, könyvtári kutatási tevékenységek végzése. Fontosságukat általában alábecsülik tudományos szempontból, s miután a nemzeti

8 <https://oduwsdl.github.io/dsa/>

9 <https://netpreserve.org/projects/bloom-filters/>

10 <https://cc.au.dk/en/warcnet/about>

szintű kutatási háttér mindenütt nagyon korlátozott mértékű Európában, ezért is vetődött fel ötletként egy olyan hálózat létrehozása, melynek segítségével egymásra találhatnak a szakemberek, s közös és magas színvonalú tevékenységeket tervezhetnek az erőforrások hatékony felhasználása révén.¹¹ Így talán sikerül elérni azt a kritikus tömeget, melynek köszönhetően a webarchívumokra irányuló kutatások széleskörű létjogosultságot nyernek.

Sajnos a járvány következtében alig sikerült személyes részvételen alapuló találkozót tartani, melyekből legalább évente kettőt terveztek a projekt időtartam alatt (döntően a londoni, a luxemburgi és a dán aarhusi egyetemek bázisán). Az online térben is hatékonyan lehet azonban szervezni a különféle tevékenységeket. Fontos cél lett volna az egymás közgyűjteményi, intézményi, illetve kutatási tevékenységeinek jobb megismerését szolgáló, maximum öt napos tanulmányutak finanszírozása, mely az ismert helyzetben szintén háttérbe szorult. A projekt időtartamát azonban sikerült kitolni az eredetileg tervezett 2020-2021-es periódustól egészen 2023 tavaszáig. A finanszírozási háttérrel a Dán Kutatási Tanács biztosítja, a koordinátori szerepet a webtörténetírás atyjának is tekinthető Niels Brügger, az aarhusi egyetem professzora látja el. A projekt működésének három legfontosabb fóruma az operatív irányítást végző szűk testület, a plenáris munkaértekezletek, illetve a szekciókban folyó tevékenységek. A projektirányító bizottság tagjai a vezető kutatók mellett a hálózat tagjaként szereplő fiatal kutató, egy informatikus, illetve egy webarchívumban dolgozó szakember. Őket a hálózat tagjai választották meg az első értekezleten. A munkaértekezleteken az általános projektmenedzsment teendők elvégzése mellett a szekciók beszámolnak a folyamatban lévő tevékenységekről, eredményekről, és keynote előadások keretében teret adnak átfogóbb, a webarchiválást esetleg csak érintőlegesen érintő témáknak, illetve olyan előadásoknak is, melyek új kutatási irányokat alapozhatnak meg. Minden értekezleten ingyenesen részt vehetnek tanácskozási joggal nem rendelkező érdeklődők is, előzetes regisztrációt követően. Közülük számosan (mint e sorok írója is, aki a második – online – értekezleten kapcsolódott be a munkába) a projekt tagjává is válhattak. A tagok számára a fizikai térben megrendezhető értekezleteken való részvétel ingyenes, a szállásköltség is biztosítva van (sajnos eddig csak egy ilyen, részben személyes részvételen alapuló hibrid értekezlet került lebonyolításra Aarhusban, 2021 őszén). 2022-ben a járvány múltával – egy nyári londoni találkozót követően – a projekt záróértekezletére ősszel Aarhusban kerül majd sor.

A projekt hat szekcióban különféle szakmai témák köré szerveződik,¹² e szekciók munkája szervesen kapcsolódik egymáshoz. Az egyes számú szekció egyrészt nemzeti, illetve nemzetközi doméneket (webtereket) vizsgál (webhelyek száma, mérete, multimédia elemek használata, a hiperlinkekkel összefogott hálózat sajátosságai). Másrészt a felhasználókra összpontosít (milyen jellemzőik vannak és milyen webhelyeket használnak). Harmadrészt pedig vizsgálat alá vonja a globális közösségi hálózatok működését és társadalmi hatásait is. A második munkacsoport a különféle kiemelt események webes megjelenésére, s az ebből adódó kutatási, elemzési lehetőségekre összpontosít. A harmadik szekció a webarchiválással összefüggő módszertani teendőkre fókuszál. A negyedik munkacsoport a webarchiválás során létrejövő és a kutatás során használható adattömeg kezelésére jött létre. Sok szó esett benne a szerzői jogi, illetve adatvédelmi dilemmákról és jó gyakorlatokról a webarchívumokhoz való hozzáférés kapcsán. Az ötödik munkacsoport voltaképpen a negyedikben vázolt kutatási adatkezelés egy speciális összetevőjére összpontosít: a webarchívumokból származó kutatási adatok szabványos megosztására. A hatodik munkacsoportban a kutatási hálózat hosszabb

11 <https://cc.au.dk/en/warcnet/the-warcnet-team>

12 <https://cc.au.dk/en/warcnet/working-groups>

távon való fenntarthatóságra koncentrálnak a WARCnet pályázat lejárata utáni időszakra. Felmérik a különböző európai uniós pályázati lehetőségeket, az egyes, már létező nemzeti, illetve nemzetközi kutatási hálózatokba történő bekapcsolódás formáit.

Az első, a harmadik és az ötödik munkacsoporthoz kapcsolódva fontos kérdésként vetődött fel, hogy miként lehetne becsatornázni a kutatói igényeket a webarchívumok szolgáltatási környezetébe, hogyan lehet azokat egyáltalán megfogalmazni, milyen adatformátum és szoftveres eszköztár szükséges a kutatásokhoz. Miként lehetne megteremteni egyfajta átjárható közös alapokkal bíró szolgáltatási és adatformátum környezetet a kutatók számára, mely egységes alapokat nyújt az egyes nemzeti webarchívumokban zajló vizsgálódásokhoz.

Bemutatásra kerültek továbbá az eseményalapú adatelemzési munka első eredményei a COVID gyűjtemény kapcsán. Elkezdtek leválogatni, hogy a felsőszintű domének szerinti eloszlásban honnan és mennyi adatot sikerült begyűjteni, továbbá vizsgálják, hogy az egyes nemzeti webterekről lementett anyagoknak milyen sajátosságai vannak, illetve összetételt, tartalmi jellemzőket tekintve milyen eltéréseket lehet tapasztalni, továbbá melyek voltak azok a főbb kulcsszavak, témakörök, melyek az archivált webtartalomban megjelentek, és milyen összetétele van a webhelyek típusa szerint a begyűjtött anyagnak általában, illetve a nemzeti webterek szintjén. A tartalomelemzésben rejlő lehetőségek bemutatására külön datathont is szerveztek.

Niels Brügger és csapata a Dán Nemzeti Könyvtár webarchívumának 2006-2016 közötti anyagára alapozva egy olyan tartalomelemzési munkát végez, ami során felmérik, hogy melyek voltak a főbb témák a dán webtérben és ezek hogyan változtak az idők során. Itt több milliónyi kulcsszó leválogatásáról és elemzéséről van szó, amihez a megfelelő erejű informatikai infrastruktúrát csak a közelmúltban sikerült felállítani.

A projekt lényeges eredményeként említhető meg a WARCnet Papers sorozat¹³, mely a kutatómunka közepette született eredmények gyors publikálásra ad lehetőséget. A sorozat első darabjai 2020 tavaszán jelentek meg és elsősorban a webarchiválással kapcsolatos kutatások perspektíváinak felvillantásával foglalkoztak. 2021-ben Niels Brügger összefoglalót készített a projekt első évének tevékenységeiről, eredményeiről. A publikációk magvát a COVID járvánnyal kapcsolatos webarchiválási tevékenységet feltáró interjúk jelentik.

Reményeink szerint a projekt lezárultával mérleget lehet majd vonni arról, hogy miképp sikerült nemzetközi szinten intézményesíteni a webarchívumokra irányuló, illetve az archivált webes anyagot forrásként használó kutatásokat, és teljes összkép állítható majd fel a projekt eredményeiről.

5. Internet Archive kereső a .hu doménról gyűjtött anyagok feltárására

Az Internet Archive egy magyar nemzeti portál kialakítására tett ajánlatot, mely tartalmazza az amerikai non-profit szervezet globális webarchívumából leválogatott összes .hu végződésű domén archivált anyagát. A portál az OSZK márkanevéhez illeszkedve kerül kialakításra, számos, a keresést segítő innovatív eszköz felhasználásának lehetőségével, melyek teljes képet igyekeznek adni arról, hogy miként jelenik meg a magyar nemzeti web domén, illetve annak egyes elemei az Internet Archive gyűjteményében.

A szolgáltatás keretében kialakítandó portál mintegy 1.2 milliárd, a .hu doménból gyűjtött URL archivált verzióját teszi elérhetővé, 1996-tól 2021 júniusáig.

13 <https://cc.au.dk/en/warcnet/warcnet-papers>

The screenshot shows the Internet Archive search interface. At the top left is the logo of the National Central Library of Hungary (Országos Széchényi Könyvtár). At the top right is the Internet Archive logo. The search bar contains the query 'kutyafajták'. Below the search bar are filters for 'webhelyek', 'képek', 'hang', 'videó', and 'PDF'. The search results are displayed in a list format, with each result showing the domain, a snippet of text, and a 'Legkorábbi verzió' link.

3. ábra: Az Internet Archive .hu domain alatti teljesszövegű keresést nyújtó portáljának találati oldala

A tömörített állomány mérete kb. 66 TB. A portál alapvető szolgáltatásai a következők: Részletes keresést lehetővé tevő funkciók, pl. kulcsszó, URL cím, dokumentumtípus (kép, videó, audió, PDF) szerint. Hozzáférés az archivált anyaghoz dedikált API segítségével, centralizált ElasticSearch klaszter felhasználásának lehetősége. Az archivált anyag megjelenítését a továbbfejlesztett Python alapú Wayback Machine eszköz biztosítja. A keresési találatok a webhelyekről és képfájlokról készült bélyegképeket is tartalmaznak. A keresési találati lista hivatkozásokat ad külön a legrégebbi, a legújabb és az összes mentés felé, valamint domén statisztikai funkciókat is kínál aggregált információkkal és MIME-típus szerinti szűréssel. A portál felületének megjelenése illeszkedni tud az OSZK honlapjához, amennyiben az intézmény igényli, és saját egyedi domén névvel is el lehet érni azt. Egy olyan verzió is rendelkezésre állhat, melyet integrálni lehet az intézmény szolgáltatási kínálatának egyéb részei mellé. A gyűjtemény egyes elemeihez történő hozzáférés korlátozására is van lehetőség, de miután a tartalom megjelenítése az Internet Archive szervereiről történik a Wayback Machine eszközzel, ezért alapértelmezésben az állomány szabad hozzáférése a kívánatos. Felhasználói analitikák és keresőlogok is az OSZK rendelkezésére állnak a szolgáltatás részeként.

Az éves fenntartáshoz és frissítéshez az alábbi szolgáltatási elemek kötődnek: a .hu doménről archivált állomány negyedévenként történő rendszeres frissítése. Az újonnan bekerülő anyag URL és kulcsszó alapú indexelése a keresőszolgáltatás számára. Webhely hosting, storage és fenntartási költségek, ideértve az összes hálózattal, sávszélességgel és karbantartással kapcsolatos díjakat is. Hozzáférés az anonimizált felhasználói és kereső analitikát érintő naplófájlokhoz. Évente 5-10 TB-ra becsült új anyag hozzáadása a gyűjteményhez. Az újonnan hozzáadott tartalmak teljesszövegű keresésének biztosítása.

Jelentős hozzáadott értékkel bír ezen felül a .hu domén anyagából összeállított zóna fájl, melynek révén az Internet Archive egyszeri listát állít össze hozzávetőleges teljességgel a .hu



domén alatti URL címekről, melyek az Internet Archive globális archívumában és kapcsolódó erőforrásai révén érhetők el. Ennek köszönhetően az OSZK jelentős mértékben tudja majd bővíteni az általános aratásba bevont webhelyek körét.

6. Epilógus

Az Országos Széchényi Könyvtárban a szervezett keretek között zajló webarchiválási tevékenység csupán 2017-ben tudott megindulni, ami nagyon jelentős hátrányt jelenthetett volna számunkra, ha nem találkozunk készséges külföldi partnerekkel, akik minden segítséget megadtak, hogy a már meglévő tapasztalatokat, jó gyakorlatokat felhasználva ki tudjuk alakítani ezen új tevékenységnek a kereteit, s be tudjunk illeszkedni a partnerintézmények által korábban megalkotott szakmai környezetbe. Nem sokkal az indulás után pedig már konstruktív módon elkezdhattünk részt vállalni a közös projektekben. Öt év múltán érdemes volt tehát áttekinteni, hogy miként alakult ki a nemzetközi kapcsolatrendszerünk, milyen együttműködési formákba tudtunk bekapcsolódni, s néhány tervet is fel tudtunk vázolni a jövőre nézve.

Irodalomjegyzék

- Brügger, Niels., Laursen, Ditte., & Schafer, Valerie (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/>
- Egense, Thomas., & Klindt Myrvoll, Anders (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/>
- IIPC. (2019). IIPC portal. Letöltés 2021. június 09; helye: <https://netpreserve.org>
- International Internet Preservation Consortium (IIPC) repository. (2020). Letöltés 2021. június 09; helye: <https://digital.library.unt.edu/explore/partners/IIPC/>
- IIPC TWG. (2019). IIPC Training Working Group. Letöltés 2021. június 09; helye: <http://netpreserve.org/about-us/working-groups/training-working-group/>
- IIPC Training Working Group. (2020). Training materials. Letöltés 2021. június 09; helye: <http://netpreserve.org/web-archiving/training-materials/>
- IIPC Training Video Case Studies. (2020). Letöltés 2020. június 09; helye: <http://netpreserve.org/web-archiving/iipc-training-video-case-studies/>
- Németh, Márton Nemzetközi körkép a webarchiválás gyakorlatáról. Könyvtári Figyelő, 63, 4.sz. (2017) 575–582. Letöltés helye: https://epa.oszk.hu/00100/00143/00349/pdf/EPA00143_konyvtari_figyelo_2017_04_575-582.pdf
- Németh, Márton. 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, Márton. Nemzetközi tanácskozás az internet megőrzéséről. Letöltés 2020. augusztus 17; helye: https://nemzetikonyvtar.blog.hu/2019/06/18/nemzetkozi_tanacskozas_az_internet_megorzeserol
- Németh, Márton. 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, Márton Webarchívum mint a tudományos kutatások tárgya. Tudományos és Műszaki Tájékoztatás, 67, 12.sz (2019), 757–765. Letöltés helye: <https://tmt.omikk.bme.hu/tmt/article/view/12804/14543>
- Németh, Márton. 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, Márton. 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, Márton. Using semantic microformats for web archiving – an initial project conception. In Katarina Tomková (Szerk.), 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. Letöltés ideje 2020. január 14; helye: <https://cloud.ulib.sk/index.php/s/JITVgF9TmRruFtC>
- Németh, Márton. 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 Márton. Webarchiválás két szakmai rendezvény tükrében. Könyv, könyvtár, könyvtáros, 28 (2019), 6.sz., 26–29.
- Németh, Márton & Drótos, László. A blended learning-based curriculum on Web archiving in the National Széchényi Library. Digital Library Perspectives, 35. 2.sz (2019), pp. 97–114. doi: 10.1108/DLP-03-2019-0012. Letöltés helye: <https://doi.org/10.1108/DLP-03-2019-0012>
- Németh, Márton & Drótos, László. Metadata Management and Future Plans to Generate Linked Open Data in the Hungarian Web Archiving Pilot Project. ITLIB, 2019(2). Letöltés helye: <https://itlib.cvtisr.sk/buxus/docs/38-metadata.pdf>
- Netlab. 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. Netlab-courses. Letöltés 2019. január 28; helye: <http://www.netlab.dk/services/courses/>
- Nielsen, J. 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
- WARCnet (2020). Letöltés 2021. június 09; helye: <https://cc.au.dk/en/warcnet/>
- The WARCnet team (2020). Letöltés 2021. június 09; helye: <https://cc.au.dk/en/warcnet/the-warcnet-team>
- WARCnet working groups (2021). Letöltés 2021. június 09: <https://cc.au.dk/en/warcnet/working-groups>
- WARCnet presentations (2021). Letöltés 2021. június 09: <https://cc.au.dk/en/warcnet/presentations>
- WARCnet papers (2021). Letöltés 2021. június 09: <https://cc.au.dk/en/warcnet/warcnet-papers>