

Az OSZK Webarchívum 2022 augusztusi hírei

Archiválás

Július végén és augusztus elején a tematikus részgyűjtemények negyedéves aratása során több mint 22 ezer webhelyről készítettünk egy újabb, a kezdőoldaltól akár 5 ugrásnyi mélységre is kiterjedő mentést (téma szerinti megoszlásuk a beszámoló végén.) Az aratások összesűritésére részben a nyári szabadságok miatt volt szükség, részben pedig a július nagy részét kitevő webtér aratás miatt kellett elcsúsztatni a szokásos ütemezést. A részletes statisztikák elkészítését is késleltette, hogy meg kellett várni, amíg a webtér archiválás utolsó fázisaként elkészült kb. 1 millió 227 ezer oldalkép, mert az azt követő aratások adatai és képei csak ezt követően kerültek a helyükre. Az aktuális webtér seed listát kereshetővé tettük a nyilvános honlapunkon a <https://webarchivum.oszk.hu/kereses-a-webter-cimlistaban/> oldalon.

A júliusban kialakított és első alkalommal le is aratott PODCAST kódnevű részgyűjteményünkhöz augusztusban is folytatódott a hangfájlok önálló lementése és elraktározása. Erre azért van szükség, mert a podcastok weboldalait archiváló robot vagy nem tudja ezeket letölteni, vagy ha van is emberi aktivitást nem igénylő letöltési lehetőség az oldalon, az archív változatban ezek a hangfájlok a legtöbb esetben nem hallgathatók meg az oldalba beágyazott lejátszóprogrammal. A hónap folyamán mintegy 400 podcast csatornáról több mint 20 ezer adást archiváltunk. A távlati terveinkben szerepel ezek cím alapján, vagy akár teljes szöveggel való kereshetővé tétele.

Az elmúlt hetekben is foglalkoztunk olyan webhelyekkel (pl. Országalbum, Ipari örökség honlap, RadixIndex honlap és RadixForum, Akadémikusválasztás 2022), melyeknek a tulajdonosai kerestek meg bennünket. Ezek archiválhatóságát vizsgáltuk, illetve a nyilvános szolgáltatáshoz szükséges szerződéskötést intéztük.

Belső szabályzat

Az OSZK megújuló Gyűjtőköri Szabályzatához még júniusban elkészítettünk egy szövegjavaslatot, ami a digitálisan születő tartalmak megőrzéséről és azon belül a webarchiválásról szól. Mivel ebben a fejezetben csak a hosszabb távon nem változó általános elvek kerültek rögzítésre, augusztus elején összeállítottunk egy ennél jóval részletesebb, sűrűbben aktualizálható, a jelenlegi munkafolyamatokat is dokumentáló anyagot „Webarchiválási munkafolyamatok szabályzata” címmel. Ebben az elvek és módszerek mellett az egyes részgyűjteménytípusok archiválási gyakorlata és azok sajátosságai is le vannak írva. Ez a dokumentum a jövőbeli webarchivátorok és a gyűjteményt használó kutatók számára lehet majd hasznos, mert kiderül belőle, hogy a projekt első 5 éves időszakában mi került a gyűjtőkörbe és hogyan archiváltuk, milyen részletességgel lett metaadatolva és ellenőrizve, készült-e róla oldalkép és statisztika, stb.

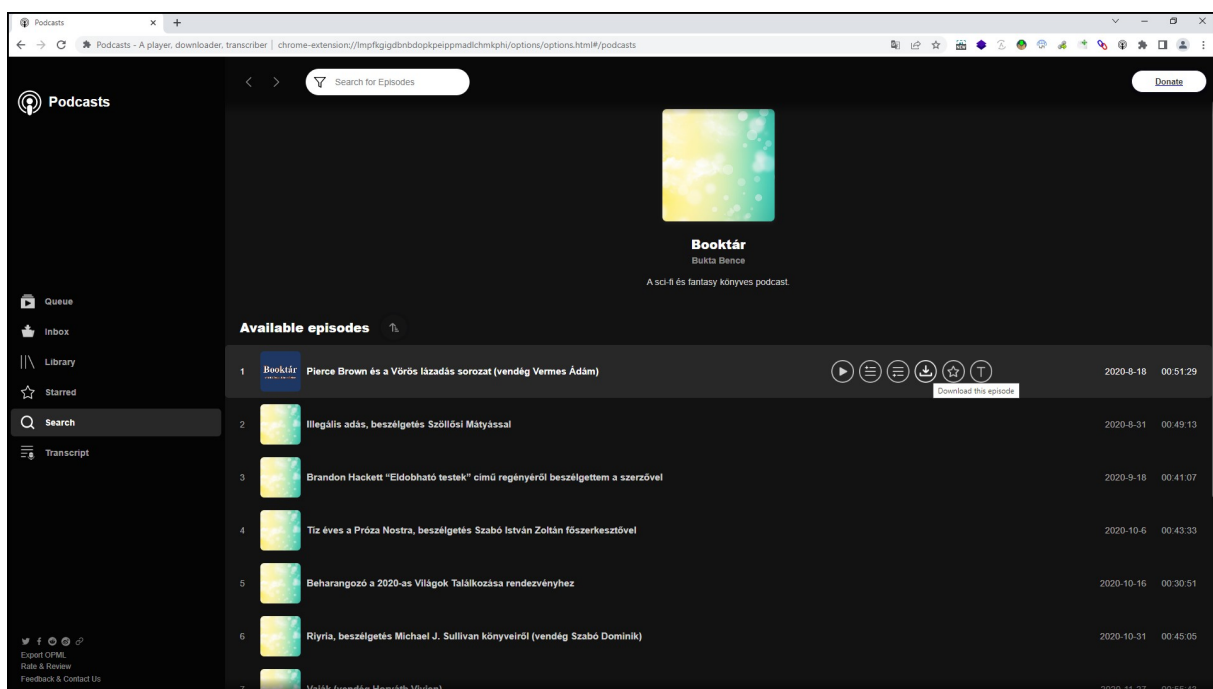
Augusztusban elkezdődött a nemzeti könyvtár adatkezelési tervének kidolgozása is, melyben a webarchívum egyik munkatársa is részt vesz.

Rendezvények

Augusztus 11-én közös rendezvényt tartott a Society of American Archivists webarchiválással és előadóművészetekkel foglalkozó két szekciója, melyen azok is részt vehettek, akik nem tagjai az SAAnak. Ennek köszönhetően Zoom-on keresztül mi is meghallgathattunk három érdekes előadást az amerikai könyvtárakban folyó munkákról. Az első a kortárs zeneszerzők webarchívumát ismertette, ami egy hibrid, vagyis hagyományos adathordozókat is tartalmazó gyűjtemény része, melyet 13

tudományos könyvtár közösen gondoz. A második esettanulmányt a Library of Congress szakemberei mutatták be, akiknek egy, a karantén időszakban elmaradó előadások miatt megszűnni kényszerülő tánc társulat webes tartalmait kellett sürgősséggel archiválniuk. A harmadik prezentáció egy analóg, digitalizált és digitális dokumentumokból álló gyűjtemény, a washingtoni városi könyvtár által gondozott punk zenei témájú archívum webes részéről szólt, külön kitérve olyan problémákra, mint a tisztázatlan gyűjtőkör és a hiányosan vagy egyáltalán nem menthető közösségi média oldalak.

Augusztus 17-én került sor az International Federation of Public History nevű szervezet konferenciáján a webarchiválás és a történettudomány témáját tárgyaló workshopra, aminek a berlini Freie Universität konferenciaközpontja adott otthont. A workshopon, melyen az OSZK webarchívumát Németh Márton képviselte, rövid bevezető előadások hangzottak el a webarchiválás elméleti és gyakorlati dimenzióiról, majd csoportmunka keretében tárgyalták meg a résztvevők, hogy milyen kisebb projekteket bonyolítanak le történészként az archivált webes tartalmakat felhasználva. Körvonalazódik egy olyan online fórum terve, ahol történészek és archivátorok tudják megosztani tapasztalataikat, illetve segíteni tudják egymás munkáját a webarchiválás terén is. A rendezvényen lehetőség nyílt a digitalizált történeti anyagok átírását segítő Transcribathon szolgáltatás kipróbálására is, amit az Europeana gondoz.



A podkasztkok archiválását bemutató videóhoz készült képernyőfotó

Ismeretterjesztés, oktatás

Az OSZK közösségi média oldalaira rövid videók készülnek, melyek a nemzeti könyvtárban folyó munka egyes részeit mutatják be. Képernyőfelvételekből mi is összeállítottunk egy mintaanyagot a podkasztkok weboldalainak és hangfájljainak archiválásáról és írtunk hozzá egy kísérőszöveget.

Augusztus elején beadtunk egy előadásjavaslatot „Special web-archiving projects at the National Széchényi Library” címmel a hosszútávú digitális megőrzéssel foglalkozó pozsonyi CDA konferenciára, melyet az ottani Egyetemi Könyvtár a járvány okozta megszakítás után idén novemberben újra megrendez.

A Könyvtári Intézet 2022. szeptember 27-30. között ismét megrendezi „Az internet archiválása mint közgyűjteményi feladat” elnevezésű tanfolyamát, melyre szeptember 13-ig lehet jelentkezni az Intézet honlapján található úrlappal. A képzés célja, hogy a résztvevők megismerjék az internetes tartalmak hosszú távú megőrzésére alkalmas technikákat, az ezzel foglalkozó külföldi projekteket és az Országos Széchényi Könyvtárban folyó webarchiválási tevékenységet.

Az elmúlt hetekben lefutott tematikus aratások

Természet- és műszaki tudományok (1488 db seed URL)
Irodalom, irodalomtudomány és -történet (1416 db seed URL)
Könyvtárak, levéltárak, múzeumok és galériák (1705 db seed URL)
Közoktatás és egyéb képzések (6544 db seed URL)
Képző-, előadó-, zene- és filmművészet (8096 db seed URL)
Történelem, hely- és családtörténet (1127 db seed URL)
Média, sajtó, műsorszórás (875 db seed URL)
Könyv- és egyéb kiadók, kereskedők (1516 db seed URL)

A tematikus aratások részletes statisztikai adatai a <https://webarchivum.oszk.hu/szelektiv-aratasok/> weblapon nézhetőek meg. A projekt hírei a <https://webarchivum.oszk.hu/a-projektrol/hirek-esemenyek/> oldalon kísérhetőek figyelemmel. Kapcsolati cím: mia@mek.oszk.hu