404 Not Found – Ki őrzi meg az internetet? Konferencia és workshop – 2022

2022/12/08
2022/11/28

Az Országos Széchényi Könyvtár Digitális Bölcsészeti Központja 2022. **december 8-án** tartja meg az internetes tartalmak archiválásával foglalkozó éves rendezvényét.

Az Országos Széchényi Könyvtár Digitális Bölcsészeti Központja 2022. december 8-án tartja meg az internetes tartalmak archiválásával foglalkozó éves rendezvényét. A hatodik „404 Not Found – Ki őrzi meg az internetet?” című konferencia és workshop újra jelenléti formában valósul meg, de az érdeklődők online is bekapcsolódhatnak az élő közvetítésbe.

A rendezvény a digitális megőrzés témáját, valamint az ahhoz kapcsolódó hazai és nemzetközi archiválási, illetve tudományos projektek ismertetését helyezi középpontba: szó lesz az OSZK-ban és

a nemzetközi szervezeteknél zajló webarchiválási tevékenységről, továbbá az előállt digitális tartalom kutatási lehetőségeiről, különös tekintettel a digitális bölcsészet szakterületeivel való kapcsolódási pontokra. A délutáni workshopon élő bemutató lesz a metaadatolásról, az egyedi webhelyek és weboldalak archiválásáról, illetve a webarchívumok keresőinek használatáról.

A TERVEZETT PROGRAM

- **10.00** ***Köszöntő***
  Rózsa Dávid főigazgató, Országos Széchényi Könyvtár

- **10.10** ***Megnyitó***
  Szóllás Péter a Könyvtári és Levéltári Főosztály vezetője, Kulturális és Innovációs Minisztérium

- **10.20** ***Az OSZK Webarchívum 2022. évi eredményei***
  Az előadás az elmúlt egy évben elvégzett munkát foglalja össze, beleértve a webarchívum bővítését, a szoftverteszteket, a tudományos és ismeretterjesztési tevékenységet, valamint a hazai és nemzetközi kapcsolatok építését. Szó lesz az újonnan létrehozott téma-, műfaj- és földrajzihely-alapú részgyűjteményekről, a fontosabb eseményekkel kapcsolatos hírek archiválásáról, a közösségi média és a podcastok mentéséről, a nyári nagy webtéraratásról, a nyilvános gyűjtemény bővítéséről és a nem publikus archívumhoz való hozzáférésről. 2022. június 1-jétől a webarchiválás a Digitális Bölcsészeti Központ szervezeti keretei között folyik tovább, így kiemelten fontossá vált a webtartalom kutatási célú hasznosítása és a tudományos munka. Az előadásban ezekről is beszámolunk, részletesebben ismertetve az IIPC-konzorcium és az idén lezáruló WARCnet projektben elért eredményeket.
  Drótos László könyvtáros, Országos Széchényi Könyvtár

- **10.45** ***Az apokalipszis archiválása: a válságok eseménytermése webarchívumokban kutatók számára***
  Ha a világméretű válságokat a digitális bölcsészettudományok szemszögéből akarjuk tanulmányozni, széles körű nemzetközi webgyűjteményekre van szükségünk, amelyek gondosan kurált és megőrzött internetes tartalmakat fűznek össze. Az International Internet Preservation Coalition (IIPC) fontos nemzetközi eseményekről született digitális webes tartalmak közös gyűjtését végzi. Ezeket az erőfeszítéseket az IIPC tartalomfejlesztési csoportjának tagjai koordinálják, akik több mint 35 ország – köztük nemzeti, egyetemi és regionális könyvtárak és archívumok – kurátorai, gyűjteményi szakemberei és webarchivátorai. A kurált tartalmakat az Archive-It-tel, az Internet Archive részével együttműködve archiválják. Az IIPC közvetlenül kutatókkal és kutatóhálózatokkal dolgozik együtt az archivált internetes tartalmak használatának és elemzésének előmozdításán. Előadásomban két kurátori együttműködésben létrehozott gyűjteményt szeretnék bemutatni: az Éghajlatváltozás gyűjteményt (2019-ben épült) és az Ukrajnai háború gyűjteményt (idén indult). Szeretném ismertetni, hogyan épült fel ez a gyűjtemény, mit tartalmaz, és milyen haszna lehet a digitális bölcsészettudományi kutatásokban.
  Dr. Teszelszky Kees kurátor, Holland Királyi Könyvtár

- **11.10** ***Adatbányászati módszer kidolgozása archivált webes tartalmakon. Az Ukrajna-projekt tanulságai***
  Az OSZK a webtérszintű és a tematikus aratások mellett eseményalapú gyűjtéseket is készít a jelentősebb kulturális, politikai és sporteseményekről. 2022. február 21. után elkezdtük gyűjteni az orosz–ukrán konfliktussal, majd később háborúval kapcsolatos híreket 75 magyarországi és határon túli portálról. A gyűjtés alapvetően a portálokon használt címkék vagy kategóriák alapján történik (ez 445 seed URL-t jelent). A mentések hetente egyszer futnak. A learatott WARC-fájlokból kibányászható adatok számos statisztikai kimutatáshoz, adatvizualizációhoz használhatók. A WARC-okból a Digitális Bölcsészeti Központ természetesnyelv-feldolgozásra alkalmas szöveget nyer ki, majd nyelvi elemzéseket végez. A

kapott adatokat összesíti és rendezi. A különböző időpontokban végrehajtott aratások anyagából szófelhők készíthetők, amelyeket megfelelő módon animálva látványosan ábrázolható a magyar hírportálok szókészletének alakulása. Az adatok idővonalon ábrázolhatók, a különböző beállításokkal számos megjelenítési lehetőség elérhető. A kidolgozott módszer alkalmas arra, hogy a gyűjtemény bármely részéből nyelvileg elemezhető szövegkorpuszokat építsünk, és azokból szövegbányászati módszerekkel adatokat nyerjünk.
Dr. Kalcsó Gyula digitálistartalom-fejlesztő, Országos Széchényi Könyvtár

- **11.35 *Tematikus digitális gyűjtemények elemzése az AVOBMAT többnyelvű kutatási eszközzel***
  Az előadás célja, hogy bemutassa az AVOBMAT (Analysis and Visualization of Bibliographic Metadata and Texts) többnyelvű kutatási eszköz működéséhez kapcsolódó munkafolyamatot és a különböző elemzőfunkciókat. A webes alkalmazás segítségével nagy mennyiségű metaadatot és szöveget lehet kinyerni és kritikusan elemezni adatvezérelt mesterséges intelligenciával és természetesnyelv-feldolgozásos technológiákkal támogatott módszerekkel és eszközökkel. Az AVOBMAT szöveg- és adatbányászati eszköz újdonságai a következők: számos nyelven képes előfeldolgozni, (szemantikusan) gazdagítani és elemezni metaadatokat és szövegeket; a beépített funkciók lehetőséget adnak a szoros és távoli olvasásra egyaránt; egy felhasználóbarát, interaktív grafikus felületen integrál metaadat- és szövegelemzéssel kapcsolatos kutatási eszközöket. A platformfüggetlen alkalmazás elsősorban olyan felhasználók számára lett kifejlesztve, akiknek nincsenek programozási ismereteik. Az egyszerűen használható felület interaktív paraméter-beállítást és vezérlést biztosít a normalizálást is támogató előfeldolgozástól az analitikai szakaszokig. A felhasználók interaktív módon kísérletezhetnek az elemzések különböző beállításaival a munkafolyamat során. Ezáltal az AVOBMAT segít felismerni a számítógépes szöveg- és adatelemzés episztemológiai kihívásait, korlátait és erősségeit, valamint kritikus módon értelmezni az alkalmazott módszereket és eredményeket.
  Dr. Péter Róbert docens, Szegedi Tudományegyetem

- **12.00 Ebédszünet**

- **13.00 *A virtuális kiállítástól az archiválásig: Karikó Katalin nyomában***
  Az előadás azt az utat mutatja be, ahogy a korábban megpendített „szegedikumok” anyagából kiemeltük a Karikó Katalin-témát és virtuális kiállítást készítettünk belőle, melynek része lett egy nemzetközi kitekintésű Karikó-weblinkgyűjtemény is. Ennek kapcsán a webtérből elkezdtük gyűjteni a lementhető és efemernek tűnő anyagokat, hogy házilagosan file-szinten archiváljuk őket. Ezután a MIA projekttel együttműködve ajánlójegyzéket kezdtünk készíteni azokról a webhelyekről, amelyeket elmentésre javaslunk egy Karikó-archívumhoz. Az előadás érinti a tervezés, kivitelezés mozzanatait, a gyűjthetőség, lementés, archiválás nehézségeit, a szerzői jogi kérdéseket és az archiválás és/vagy szolgáltatás, redundancia és teljesség mindnyájunkat gyötrő dilemmáit is.
  Dr. Kokas Károly informatikai főigazgató-helyettes, Szegedi Tudományegyetem Klebelsberg Kuno Könyvtára

- **13.25 *Hogy kerül egy oklevél a webarchívumba? Aranybulla-emlékév-archívum kialakítása a Vörösmarty Mihály Könyvtárban***
  #webarchiválás #vmk #Aranybulla800 #Székesfehérvár
  #pilotprojekt #kihívás #kitartás #csapatmunka #helyismeret #feldolgozócsoport #olvasószolgálat
  #fehérváriak #kutatók #diákok #pedagógusok
  #tudományos #ismeretterjesztő #médiatörténet #mindennapitörténetünk #források #hírek #cikkek
  #munkafolyamat #nehézségek #fókuszpontok #tapasztalatok
  Horváth Adrienn igazgatóhelyettes, Vörösmarty Mihály Könyvtár

- **13.50 Szünet**

- **14.00–16.00 *Workshop könyvtárak számára***
  Drótos László – Visky Ákos László, Országos Széchényi Könyvtár

Ingyenes regisztráció [1] az OSZK honlapján

A korábbi „404 Not Found” rendezvények anyaga az OSZK Webarchívum honlapján [2]

Kapcsolati e-posta-cím: mia[kukac]mek.oszk.hu

A rendezvény plakátja [3]
2022/12/04 - 12:06

Forrás webcím: https://oszk.hu/rendezvenyek/404-not-found-ki-orzi-meg-az-internetet-2022

Hivatkozások:
[1] https://www.oszk.hu/404-not-found-2022-regisztracio
[2] https://webarchivum.oszk.hu/404-not-found-workshop/.
[3] https://oszk.hu/sites/default/files/Plakat_404-Konferencia_2022-HU.pdf
[4] https://oszk.hu/category/foszotar-es-pozicionalo/hirek
[5] https://oszk.hu/category/foszotar-es-pozicionalo/rendezvenyek