A KDS keretében megvalósítandó tematikus webarchívum és módszertani ajánlás munkaterve

Az OSZK 2019-es KDS pályázatának részeként kialakításra kerül egy, a közoktatásban is felhasználható tematikus digitális gyűjtemény, valamint kidolgozzuk a webtartalmak közösségi gyűjtésének módszertani ajánlását.

A gyűjtemény témája a „II. Rákóczi Ferenc emlékév", mely az Országgyűlés 32/2018-as határozata alapján 2019. július 8-tól 2020. szeptember 17-ig tart. Az emlékévvel kapcsolatos események „internetes lenyomatait" gyűjtjük össze különféle webarchiváló technikákkal, valamint lementünk további, Rákóczival és a szabadságharccal foglalkozó weboldalakat vagy teljes webhelyeket. Ezeket a dokumentumokat metaadatokkal látjuk el és böngészhetővé, illetve kereshetővé tesszük egy erre a célra létrehozott honlapon. A keresőbe felvesszük az E-könyvtári Szolgáltatások Osztály kezelésében lévő egyéb nyilvános digitális gyűjtemények ilyen témájú dokumentumait is, melyek között szép számmal vannak szintén az internetről gyűjtött, valamint digitalizált könyvek, képek, periodikákban megjelent cikkek és tanulmányok. Így ez a szolgáltatás amellett, hogy bemutatja a webarchiválás lehetőségeit és korlátait, arra is jó példa lesz, hogy hogyan lehet integrálni egy könyvtár „hagyományos" digitális dokumentumgyűjteményét a – nálunk még újszerűnek számító – webarchívummal. A projekt keretében kifejlesztett informatikai megoldások az OKP rendszer igényeinek figyelembe vételével készülnek el.

A főbb részfeladatok:

- az archiválandó weboldalak/webhelyek kiválasztása (kb. 300 tétel);
- szerződéskötés a tartalomgazdákkal az archivált példány nyilvános szolgáltatására (ahol nem sikerül engedélyt szerezni, ott csak a metaadatok és egy kis méretű oldalkép lesz nyilvánosan megtekinthető);
- a kiválasztott webes tartalmak archiválása és minőségellenőrzése (komolyabb hibák/hiányok esetén újramentés más technológiával);
- oldalképek készítése az eredeti külalak pontos megőrzése és a minőségellenőrzés céljából;
- a mentett tartalmak metadatolása egy egyszerűsített XML séma alapján;
- a MEK, a DKA és az EPA szolgáltatásokból Rákóczi-témájú dokumentumok válogatása (gyűjteményenként 200-300 tétel);
- a kiválasztott dokumentumok metaadatainak átalakítása a webes gyűjtemény XML sémájának megfelelően;
- metaadat-szintű kereső- és böngésző alkalmazás fejlesztése, mellyel a webarchívum anyaga külön és együtt is láttatható a másik három gyűjtemény anyagával;
- teljes szövegű és képkereső beüzemelése az archivált webes tartalmakra;
- szolgáltatófelület kialakítása, ahol a felhasználók a gyűjtemény használata mellett tájékoztatást kapnak a webarchiválási technológiákról és az OSZK webarchiváló projektjéről, és ők maguk is ajánlhatnak további, megőrzésre érdemes webes tartalmakat (a fontosabb tudnivalók a magyar mellett további 5-6 nyelven is olvashatók lesznek);
- módszertani ajánlások készítése, melyek alapján más közgyűjtemények vagy egyéb szervezetek is be tudnak kapcsolódni a magyar webtér oktatási, kulturális és tudományos tartalmainak hosszú távú megőrzésébe (résztémák: válogatás, engedélyeztetés, archiválás, oldalkép készítés, minőségellenőrzés, metaadatolás);
- projektirányítás, adminisztrálás, helpdesk, eredmények disszeminációja (kiemelten a közösségi médiában és más internetes fórumokon).

Készült: 2019.10.04.