

Az Europeana Travel projekt megvalósítása a Debreceni Egyetem Egyetemi és Nemzeti Könyvtárban

Köpösi Zsuzsa

Debreceni Egyetem Egyetemi és Nemzeti Könyvtár

Networkshop, 2011. április 28.

Az Europeana Travel projekttel a Debreceni Egyetem Egyetemi és Nemzeti Könyvtár (DEENK) első ízben vett részt európai szintű digitalizálási együttműködésben, amely jó lehetőséget kínált új tapasztalatok megszerzésére és új technikai megoldások kipróbálására, alkalmazására. A DEENK számára a tapasztalatszerzésen kívül az is fontos cél volt, hogy a projekt megvalósításával a könyvtár digitális állományának jelentős része elérhetővé váljon az Europeana portálon (<http://www.europeana.eu/portal/>) keresztül is, amely a projekt időpontjában már több mint 1.500 intézmény közel 15 millió digitális dokumentumát tartalmazta.

Az Europeana Travel projekt 2009 májusában indult 16 ország 17 – egyetemi, kutatóintézeti és nemzeti – könyvtárának részvételével. A kétéves projektben Magyarországról egyedül a DEENK vett részt. Az Europeana Travel projekt egyik fontos célja a résztvevő könyvtárak állományában megtalálható, az utazás és turisztika témakörébe tartozó forrásdokumentumok (térképek, könyvek, fotók, képeslapok, folyóiratok) digitalizálása, és az Europeana portálon keresztül bárki számára való elérhetővé tétele. A

projekt másik fontos célkitűzése új intézmények digitális állományának bekapcsolása a portál adatbázisába. Ennek eléréséhez fontossá vált az úgynevezett LIBER aggregátor felállítása, amelyen keresztül a projektben résztvevő – az Europeana által eddig még nem aratott – LIBER tagkönyvtárak digitális állományai váltak elérhetővé az Europeana felhasználói számára.

A DEENK a projekt keretében könyveket (160 cím, 46.000 oldal), térképeket (200 cím) és folyóiratot (1 cím, 15.000 oldal) digitalizált. A digitalizálást 600 DPI felbontásban, TIFF fájlok létrehozásával végeztük el. A szöveges dokumentumokon javítás nélküli szövegfelismertetést végeztünk, és kétrétegű PDF fájlokat hoztunk létre. A projekt keretében elkészült digitális dokumentumok online szolgáltatása a DEENK digitális könyvtárán keresztül biztosított. A DEENK digitális könyvtára a Debreceni Egyetem Elektronikus Archívumában (DEA-ban) található. A DEA érdekessége így, hogy nem csak a Debreceni Egyetem intézményi repozitóriumaként működik, hanem a DEENK digitális állományát szolgáltatva digitális könyvtárként is funkcionál. A DEA működtetését, karbantartását és menedzselését a DEENK munkatársai végzik. A DEA a DSpace nevű keretprogram használatával működik, amely OAI-PMH kompatibilis és a Dublin Core használatát támogatja.

A DEENK-ben alkalmazott munkafolyamat során a digitalizált forrásdokumentumokról bibliográfiai leírások készültek MARC formátumban a DEENK – Corvina rendszerben működtetett – online katalógusában. Az így létrehozott MARC rekordok felhasználásával, a digitalizálás során elkészített TIFF, és PDF fájlokat automatikus konverzióval töltöttük be a DEA-ba. Az automatikus konverzió során a DEA-ban létrejöttek a megfelelő metaadatrekordok, amelyek az adott digitális dokumentumok – MARC rekordból konvertált – DC alapú metaadatait, valamint a forrásdokumentumról elkészített, felhasználók által is látható szolgáltatási PDF fájlt és a rejtett, archiválási TIFF fájlokat tartalmazzák. Ehhez össze kellett állítanunk mindhárom dokumentumtípusra vonatkozóan a MARC–DC mezők megfeleltetését tartalmazó konverziós táblázatokat.

A munkafolyamat második része során a DEENK digitális könyvtárában már szolgáltatott digitális állományt kellett az Europeana számára arathatóvá tenni, ezáltal

biztosítva, hogy az Europeana portálon végzett keresések során a projekt keretében elkészített digitális dokumentumok elérhetővé váljanak. Ehhez a LIBER aggregátor számára megfelelő formátumú adatok OAI-PMH protokoll szerinti szolgáltatását kellett megoldani. Az adatokat az Europeana által használt ESE (Europeana Semantic Elements) szabványnak megfelelően kellett szolgáltatunk. Ehhez el kellett készítenünk a DEA-ban használt DC adatok és az Europeana által használt ESE megfeleltetéseit tartalmazó táblázatokat. Valójában a MARC–DC és a DC–ESE megfeleltetéseket együtt, egy táblázatban végeztük el. Az Europeana által a LIBER aggregátoron keresztül végzett aratás során kerültek be a DEENK digitális állományából a projekt keretében elkészített digitális dokumentumok leíró adatai, és a dokumentumokra mutató közvetlen linkek az Europeana adatbázisába.

A konverzió és az aggregálás előkészítéséhez nem csak a MARC-DC-ESE táblázatok elkészítésére volt szükség, hanem egyéb fontos döntések átgondolására, és meghozatalára is. Így el kellett dönteni, hogy a DC elemek közül melyek jelenjenek meg a DEA áttekintő metaadatrekord-megjelenítésében, illetve hogy milyen megnevezéssel jelenjenek meg a felhasználók számára. Mindeközben fontos volt tudnunk azt, hogy a teljes projektben egymással párhuzamosan több igénynek is meg kell felelnünk, hiszen az egyes felhasználók számára is minél megfelelőbb szolgáltatást szeretnénk nyújtani, miközben az intézményünk érdekeit is szem előtt kellett tartanunk és az Europeana elvárásainak is meg kellett felelnünk. Ezen túl már az előkészítés során szembe kellett néznünk az egyéb, egyedi problémák felmerülésével, amelyek közül egy érdekes probléma részleteit az alábbiakban megtekinthetjük.

Az ESE rendszere nem csak az eredeti 15 DC metaadatelemet, hanem az Europeana igényeinek megfelelő további 12 elemet is tartalmazza.

| | | |
|-----------|-------------|--|
| | | |
| DC | title | alternative |
| DC | creator | |
| DC | subject | |
| DC | description | tableOfContents |
| DC | publisher | |
| DC | contributor | |
| DC | date | created; issued |
| DC | type | |
| DC | format | extent; medium |
| DC | identifier | |
| DC | source | |
| DC | language | |
| DC | relation | isVersionOf; hasVersion; isReplacedBy; replaces; isRequiredBy; requires; isPartOf; hasPart; isReferencedBy; references; isFormatOf; hasFormat; conformsTo isShownBy; isShownAt |
| Europeana | | |
| DC | coverage | spatial; temporal |
| DC | rights | |
| DC terms | provenance | |
| Europeana | userTag | |
| Europeana | unstored | |
| Europeana | object | |
| Europeana | language | |
| Europeana | type | |
| Europeana | uri | |
| Europeana | year | |
| Europeana | hasObject | |
| Europeana | country | |

Az ESE rendszerben a type és a language elemek mind a DC, mind az Europeana elemek között megtalálhatók. Nézzük meg részletesebben a language elem esetét. A dc.language elem értéke a digitális dokumentum nyelvének a megfelelő kódja. Az europeana.language elemnek viszont az adott digitális dokumentumot szolgáltató intézmény

székhelyének helyet adó ország hivatalos nyelvét kell tartalmaznia. Ebből adódóan, ha az Europeana portálon a portál keresőfelületén a találatokat „Hungarian” (magyar) nyelvre szűkítjük, a találati listában olyan tétel is meg fog jelenni, amely más (nem magyar) nyelvű dokumentum adatait tartalmazza. Erre az esetre példa Gerhard vom Rath: Siebenbürgen című művének tétele az Europeana portálon, amelynek dc.language eleme a „hu” értéket, az europeana.language eleme pedig a „ger” értéket tartalmazza.

A projekt csak a szerzői jogilag már nem védett dokumentumok digitalizálását támogatta, így a dokumentumok kiválogatásához kialakított szempontrendszerben ez a kitétel kiemelt helyen szerepelt. Az Europeana portálon a digitális anyagok leíró adatai (metaadatai) válnak elérhetővé, és kereshetővé, maguk a digitális dokumentumok az adott tartalomszolgáltató intézmény saját digitális könyvtárában tárolódnak.

A könyvek esetében a munkafolyamatok figyelemmel kísérésére új alkalmazást próbáltunk ki – a Dotproject projektmenedzsmet szoftver felhasználásával történt az egyes dokumentumokon végrehajtott munkafázisok követése. Ez pontos, naprakész, megbízható adminisztrációs módnak bizonyult, hiszen minden digitalizálási projektben fontos a dokumentumok folyamatos védelme, illetve pontos nyomonkövetése is.

Az projekt keretében elkészített digitális anyag elérhető a DEENK digitális könyvtárában, a DEA-ban <http://ganymedes.lib.unideb.hu:8080/dea/handle/2437/1>, illetve az Europeana-n keresztül is <http://www.europeana.eu/portal/>. Néhány példán keresztül érdemes megtekinteni az digitális dokumentumok és leíró adataik megjelenési formáit a különböző elérési lehetőségeken keresztül.