

Tartományi Könyvtár tapasztalatai alapján még várat magára.

/LAMERZ, Lennart: E-Books in der Wissenschaft – Fluch oder Segen? Eine Untersuchung der Nut-

zungsstatistiken der Universitäts- und Landesbibliothek Düsseldorf im Bereich E-Books = Informaton. Wissenschaft & Praxis, 63. köt. 4. sz. 2012. p. 233-240./

(Némethné Szivi Zsófia)

A Szlovén Nemzeti és Egyetemi Könyvtár részvétele az IMPACT projektben

Hasznára válik-e a Szlovén Nemzeti és Egyetemi Könyvtár (*National and University Library of Slovenia = NUK*) számára az IMPACT projektben való részvétel? A cikk, mielőtt megadná a választ a kérdésre, bemutatja a NUK tevékenységét, felhasználóit, s magát a projektet is, összehasonlítva azt más, szintén uniós finanszírozású projektekkel, amelyekben a könyvtár részt vesz. Általánosabban fogalmazva, a szerző a következőkre keresi a választ:

- hasznos-e a nemzeti könyvtárak részvétele a különböző nemzetközi kutatási projektekben, s ha igen, miért;
- a tömeges digitalizálással kapcsolatos kihívások, problémák kezelhetők-e intézményi vagy nemzeti szinten, avagy csak nemzetközi összefogás adhat rájuk megoldást.

Az IMPACT projekt

Az IMPACT projekt 2008. január 1-jén indult el a Holland Nemzeti Könyvtár vezetésével, a 7. Kutatás-fejlesztési Keretprogram részeként. Célja az volt, hogy automatikus szövegfelismerő rendszerek kidolgozásával és az OCR technológia továbbfejlesztésével javítsa a hozzáférést a tömeges digitalizálási projektek keretében létrejövő, az európai kulturális örökséget feltáró digitális gyűjteményekhez. A projekt négy alprojektje két fázisban zajlott le. Míg az első szakaszban 15 intézmény vett részt (7 könyvtár, 6 kutatóintézet és egyetem, 2 vállalkozás), jelenleg 13 ország 26 intézménye vesz részt a fejlesztésekben (12 könyvtár, 12 kutatóintézet és egyetem, 2 vállalkozás). A NUK a második szakaszban, 2010 elején csatlakozott a projekthez.

A projekt fő célkitűzései a következők: az OCR technológia továbbfejlesztésével javítani a történeti művekhez való digitális hozzáférést; a tömeges digitalizálás során alkalmazott nemzeti és nemzetközi megoldások elemzése; ajánlások megfogalmazása a *best practice* elve alapján; egy kompetenciaközpont (Centre of Competence) létrehozása a projekt végére a fejlesztések nemzetközi szinten történő elérhetővé tétele érdekében. (Az OCR fejlesztések elsősorban a nyelvi sajátosságokból származó hibák, a dokumentumok nyomtatási minőségéből adódó hibák, illetve a rossz minőségű digitalizálásnak köszönhető hibák kiküszöbölését célozzák meg.)

A projektben európai nemzeti és egyetemi könyvtárak sora képviseli a felhasználók igényeit és szükségleteit, de emellett más fontos feladatok ellátása is a partnerkönyvtárakra vár (pl. projektmenedzsment; tartalomszolgáltatóként a tudományos kutatáshoz szükséges dokumentumok biztosítása; a projekteredmények tesztelése, értékelése és promóciója). A fejlesztéseket a tagkönyvtárak vagy partnereik által digitalizált szövegeken tesztelték; az OCR-ezett szövegeket manuálisan javították, majd újra ellenőrizték. A tartalomszolgáltatási feladatokon túl néhány könyvtár az egyes munkacsomagok vezetőjeként is tevékenykedett, míg a második fázisban csatlakozók kisebb feladatokat oldottak meg. A könyvtárak hasznos információkkal szolgáltak a projekt számára a tömeges digitalizálás során felmerülő problémákkal kapcsolatban is. Ezek az információk jelentősen befolyásolták a szoftverfejlesztések irányát és tartalmát. A NUK a szláv nyelveket beszélő felhasználói közösség speciális igényeit is meg tudta jeleníteni a projekt végrehajtása során.

A Szlovén Nemzeti és Egyetemi Könyvtár

A Szlovén Nemzeti és Egyetemi Könyvtár évente mintegy 200 000, a szolgáltatásokat helyben használó és több mint egymillió távoli felhasználót lát el. A távoli felhasználók leggyakrabban a könyvtár portáljáról (*dLib.si*) elérhető információs forrásokat keresik: nemzetközi és szlovén időszaki kiadványokat, s más online adatbázisokat. A portál felhasználói jellemzően egyetemisták (71,30%), átlagemberek (21,50%), egyetemi oktatók és kuta-

tók (3,10%), diákok (1,80%), illetve külföldiek (2,30%). E felsorolás jól reprezentálja egy modern könyvtár felhasználói körét, akiknek a nyomtatott dokumentumok mellett szükségük van a digitális forrásokra is, s ezeket otthonról, munkahelyükről is szeretnék elérni – a legkülönbözőbb technikai eszközök segítségével. (Lényegében nem különböznek a könyvtárat személyesen használó felhasználói csoportoktól.) A Google Analytics segítségével (amely meg tudja különböztetni a tényleges használókat a keresőrobotoktól) a könyvtár több mint 2,8 millió látogatást regisztrált a portálon, javarészt Szlovéniából (93%).

A *dLib.si* oldalon elérhető források között állományvédelmi céllal digitalizált műveket, illetve a szlovén kulturális örökség megőrzésének céljával digitalizált dokumentumokat egyaránt találunk. 2011 júniusában a portál több mint 4 millió, szabad hozzáférésű szkennelt dokumentumot tett elérhetővé.

A portált a felhasználók általában hagyományos webes keresőmotorokkal használják, s teljes szövegű dokumentumokat keresnek rajta. A gyenge minőségben OCR-ezett szövegek azonban jelentősen megnehezítik számukra a releváns találatokhoz való hozzájutást. A statisztikák azt mutatják, hogy szinte egyelő arányban választják a dokumentumok megtekintésére a HTML és a PDF formátumokat, annak ellenére is, hogy az előbbit az oldaltörések miatt sokkal nehezebb olvasni. A HTML formátum ugyanakkor megkönnyíti az egyes szövegrészek másolását, módosítását és további használatát. A PDF verzió ezzel szemben visszaadja az eredeti dokumentum külsejét, s könnyű letölteni, illetve a számítógépre menteni. Az 1850 előtti, szlovén nyelvű dokumentumok esetében az OCR-ezett szöveg azonban nagyon gyenge minőségű, s a HTML verzió nemcsak egy felismerhetetlen karakterhalmazt mutat, hanem rengeteg hibát és furcsa jelet is, betűk helyett számokat stb. Ezeknél a dokumentumoknál gyakorlatilag lehetetlen a teljes szövegű keresés, illetve a szöveg további használata.

Az ilyesféle komplex problémák megoldása csak különböző szakértők közös munkájával lehetséges – ehhez nyújthatnak nagy segítséget az EU-projektek, amelyekben a NUK már 2001-től részt vesz, s amelyek között sok hasonlóság és sok különbség is felfedezhető.

A DRIVER II projekt esetében a könyvtár hasonló finanszírozási struktúrában (7. K+F Keretprogram) tevékenykedett, mint az IMPACT esetében, s a

konzorciumban is könyvtárak és egyetemek együtt vettek részt. A projekt eredményeképpen egy olyan, a nyílt hálózatok elvén működő szolgáltatás jött létre, amely könnyen adaptálható bármely régióra. A *Nyílt Hozzáférésű Adattárak Szövetségének (Confederation of Open Access Repositories – COAR)* a világ minden tájáról van tagja. Az IMPACT hálózatát egyszer nyitották csak meg a projektzárás után: az *IMPACT Centre of Competence (CoC)* 2011. októberi megnyitásakor kiindulási alapot biztosított a hálózat további működéséhez.

A keretprogram mellett az eContentPlus, illetve később a CIP ICT PSP program biztosított fejlesztési forrásokat – egy más támogatási struktúrában: egyrészt magának az *Európai Digitális Könyvtárnak (Europeana)* a létrehozását támogatta, másrészt különböző alprojektekhez kínált pénzügyi segítséget (Europeana Travel, Europeana Local). A NUK mindhárom projektben részt vett.

Az Europeana nyitott hálózatként működik, s folyamatosan ösztönzi is az új tagokat (intézményeket, szervezeteket, projekteket) a csatlakozásra. Ebben a hálózatban a Szlovén Nemzeti és Egyetemi Könyvtár egyrészt nemzeti aggregátorként tevékenykedik, másrészt pedig elsődleges kontaktként szolgál a szlovén tartalomszolgáltatók számára Europeana ügyekben. Az Europeana-tagsággal a NUK átvette és közvetíti a hálózatban használt szabványokat, munkafolyamatokat is; ez biztosítja a *dLib.si* oldalon elérhető metarekordok magas minőségét is.

A digitalizálási projektek mellett a könyvtár részt vesz az Európai Unió Kultúra 2007–2013 programjából 50%-ban finanszírozott *EOD – eBooks on Demand (e-könyvek kérésre)* elnevezésű projektben is. Az EOD hálózat a kontinens könyvtárainak és levéltárainak folyamatosan gyarapodó partnerségi köre, amely lehetővé teszi, hogy a felhasználók aktívan alakíthassák a közgyűjtemények digitalizálási terveit. Az EOD kiváló visszajelzést ad arról, hogy mire van szüksége a használóknak, s befolyásolni tudja, hogy a könyvtárak – így a NUK is – milyen prioritások mentén alakítsák digitalizálási programjaikat.

Az EOD és az IMPACT között érdekes kapcsolat figyelhető meg: a felhasználói kérésre történő digitalizálás során a szakemberek azokat a munkafolyamatokat és megoldásokat alkalmazzák a digitális objektumok minőségének javítására, amelyeket az IMPACT keretében fejlesztettek ki. A digitalizált

könyveket a könyvtárhasználók nyomtatott formában és OCR-ezett PDF-ben is megkaphatják.

A nemzetközi projektekben való részvétel előnyei és kihívásai

Előnyök

- Új ismeretek és a legjobb gyakorlatok elsajátítása; a projektek révén a könyvtárak elméleti tudásukat gyakorlati megoldásokká alakíthatják. A NUK például a DRIVER II projekt révén megismerte és gyakorlatban is alkalmazni kezdte az új metaadatszabványokat és tartalomcsere-protokollokat. Az IMPACT révén minden addiginál gyakorlatiasabb ismeretre tett szert a digitalizálásban, s megtanulta az együttműködésből adódó tapasztalatokat is a maga javára fordítani.
- A nemzetközi szabványok kifejlesztésében végzett kollaboratív munka ugyancsak számos előnnyel jár az egyes intézmények számára: megtanulják képviselni és kifejezni saját gondolataikat, tapasztalataikat és ismereteiket, ugyanakkor lehetőségük van tanulni, a mások által kitalált megoldásokat adaptálni, esetleg továbbfejleszteni. Ez történt a szlovén nemzeti könyvtárral is az Europeanban és az EOD-ban. Ráadásul az új szabványok bevezetése és alkalmazása sokszor más projektekben való részvételnek is előfeltétele.
- Nemzetközi hálózatokban való együttműködés és integrálódás. A szélesebb nemzetközi szakmai közösségbe való bekapcsolódással lehetővé vált a NUK számára, hogy más, sokkal objektívebb nézőpontból szemlélje a digitalizálással kapcsolatos kihívásokat, problémákat.
- A projektmunka általában munkacsomagokban zajlik, s a partnerek megosztják a feladatokat egymás között. A projektek sikeréhez elengedhetetlen a munkatársak aktív részvétele és a csapatmunka – ahogy ezt a NUK is megtanulhatta az IMPACT projektből, ahol a legkülönbözőbb szakterületek specialistái dolgoztak együtt.
- A kutatók és intézmények nemzetközi hálózatához való csatlakozás erősíti az adott könyvtár nemzetközi szakmai közösségben elfoglalt helyét, amely egy olyan kis állam nemzeti könyvtárának különösen fontos, mint Szlovénia.
- A projektek révén erősödik az intézmény nemzeti szintű együttműködése is; a NUK számos új könyvtárral, egyetemmel, kutatóintézettel létesített szakmai kapcsolatokat az IMPACT keretében.
- Végül, de nem utolsósorban, előny az új források lehetősége. Az EU-projektek olyan pénzügyi tá-

mogatásokhoz juttatják a könyvtárakat, amelyek nélkül talán nem is tudnának új fejlesztésekbe kezdeni.

Kihívások

Az előnyök mellett a projektek megjelenése a könyvtárak életében számos kihívással is jár. A NUK tapasztalatai alapján ezek a következők:

- Fenntarthatóság – a legtöbb probléma ebből adódik, vagyis a projekteredmények bevezetéséből, fenntartásából és a projektmunkatársak továbbfoglalkoztatásából. A kettő pedig szorosan összekapcsolódik: ha nincs pénz a munkatársak foglalkoztatására, nincs, aki megfelelő tudással és gyakorlattal rendelkezik a projekt eredményeit, tapasztalatait illetően.
- Finanszírozási problémák, azaz a projekt végrehajtására fordítandó pénzek – az uniós projekteknél az önköltség – biztosítása. A források konkrét célokhoz kötöttek és limitáltak. Sokszor nehéz megjósolni, hogy egy konkrét feladat elvégzéséhez mennyi idő és emberi erőfeszítés szükséges, mekkora annak költsége, s milyen további költségek merülnek fel a projekt végrehajtása során (különösen több évig tartó projektek esetében).

Ahogy a fenti listából is látszik, a nemzetközi kutatási projektekben való részvételnek több előnye, mint hátránya van. Az IMPACT kiszélesítette a Szlovén Nemzeti és Egyetemi Könyvtár látókörét a digitalizálással kapcsolatos problémák felismerésében és megoldásában. A megszerzett tapasztalatokat és tudást a könyvtár meg tudja osztani nemzeti szinten is, s képes lesz arra, hogy igazi digitalizálási kompetenciaközpontként szolgáljon az ország többi közgyűjteménye számára.

Nemzeti könyvtárként a NUK felelősségi körébe tartozik a szlovén kulturális örökség (nyomtatott és digitális) megőrzése, illetve gyors és ingyenes elérhetővé tétele. Egyetemi könyvtárként pedig a fő hangsúlyt a kiváló minőségű szolgáltatásokra és a digitális források biztosítására kell helyeznie. Az IMPACT-hez hasonló projektekben való részvétel mindkét célkitűzést támogatja. Konkrétabban: az IMPACT révén több mint 5000 digitalizált történeti dokumentum (XVIII. sz. második, illetve a XIX. sz. első feléből) ellenőrzése és javítása történt meg. A most már közel 100%-os pontosságú OCR-ezett dokumentumok könnyen hozzáférhető és jól kereshető információs forrást kínálnak a felhasználók számára. A projekt további eredménye egy, a történeti szlovén nyelv szavait és sajátosságait feltá-

ró, rögzítő lexikon, amely hamarosan a *dLib.si* portálon elérhető digitális könyvtár részévé is válik, lehetővé téve a legmodernebb keresési technikák alkalmazását több mint 200 000 digitalizált történeti dokumentumban. Az együttműködés azt is megmutatta, hogy a tömeges digitalizálás problémái nemzetközi szinten sokkal könnyebben megoldhatók, s a legjobb gyakorlatok alkalmazása

mindig elvezet az eredmények maximalizálásához – költséghatékonysági szempontból is.

VODOPIVEC, Ines: The Role of the National and University Library of Slovenia in a Multinational Research Project (IMPACT): a Case Study. = Liber Quarterly. 21. köt. 3–4. sz. 2012./

(Kovácsné Koreny Ágnes)

Többletérték hozzáadása disszertációkat tartalmazó intézményi repozitóriumokhoz

Az intézményi repozitóriumokat az intézmény saját tudományos publikációinak összegyűjtésére, archiválására és közzétételére hozzák létre. Fő jellemzőjük, hogy nagyon sokfélék lehetnek a funkcióik, szolgáltatásaik, működtetési politikájuk stb. tekintetében, két dologban azonban mindig megegyeznek: teljes szöveghez adnak hozzáférést és megfelelnek a nyílt hozzáférés mozgalom (open access) által megkívánt OAI protokollnak, vagyis tartalmuk „aratható”.

Jelen cikk a disszertációkat és téziseket tartalmazó repozitóriumokat vizsgálja. Az *OpenDOAR (Directory of Open Access Repositories)* jegyzékében a repozitóriumok mintegy felénél, 1100 helyen jelzik disszertációk gyűjtését, a másik nagy nemzetközi nyilvántartó, a *ROAR (Registry of Open Access Repositories)* pedig több mint 200 olyan szolgáltatást jelez, ahol 100%-ban disszertációkat gyűjtnek. A *Tézisek és Disszertációk Digitális Könyvtári Hálózatának (NDLTD = Networked Digital Library of Theses and Dissertations)* közös katalógusa több mint egymillió tézist és disszertációt tartalmaz, melyek döntő többsége intézményi repozitóriumokban található. Ez jó hír, de azért a kérdés felvetődik: hogyan lehet a repozitóriumok hozzáférhetőségét, hatását, elfogadottságát növelni. Nem elég ugyanis a tartalmakat feltölteni a szerverekre, további értéknövelő funkciókat, szolgáltatásokat kell hozzáadni a repozitóriumokhoz.

A szerző módszere az volt, hogy 13 disszertációt tartalmazó repozitóriumokkal foglalkozó nemzeti és nemzetközi tudományos konferencia előadásait nézte végig, összesen 54 írást, melyeket a bibliográfia a cikk végén fel is sorol. Ezeket összegezve az értéknövelés öt lehetséges útját nevezi meg és részletezi őket.

1. Minőség

Egy jó repozitórium a kiválasztás feltételeként nemcsak szabványokat és kritériumokat nevez meg, hanem kommunikálja és elő is mozdítja ezt a szerkesztői politikát. A disszertációkat zsűri ellenőrzi, preprinteket és ellenőrizetlen saját archiválást nem engedélyez. A disszertációk vagy végighaladnak az intézmény által szabályozott helyi doktori eljárás elfogadási folyamatán vagy az archiválási folyamat tartalmaz egy formális intézményi validálási pontot. Ugyanez vonatkozik a régebbi, csak nyomtatott formában létező disszertációk digitalizálása után történő bekerülésre.

2. Jó és szabványos metaadatok

A jó minőségű és sok metaadattal ellátott rekordok hozzájárulnak a jó kereshetőséghez és tudományometriai elemzések végzéséhez. Különösen fontosak a metaadat-szabványok a kis és közepes méretű intézmények számára és a DART-Europe-hoz hasonló hálózatokban. A DART az OAI Dublin Core egyszerűsített formáját alkalmazza. Akármelyik metaadatsémát is használja azonban a repozitórium, annak rugalmasnak és más formátummá könnyen konvertálhatónak kell lennie.

3. Formátum

Az adattár formátuma legyen jól kereshető, nyitott és alkalmas a hosszú távú megőrzés céljaira. Ezen túlmenően biztosítsa az intelligens kiaknázhatóságot, vagyis az adatokat lehessen importálni és újra felhasználni. Tárolásra elfogadott formátumok az MS Office Word és az Adobe PDF, de vannak olyan repozitóriumok, amelyek elfogadják az XML, LaTeX, Postscript vagy RDF formátumokat is. Retrospektív digitalizálás után szükséges az OCR-ezés (a szöve-