

Az adatbázis-egyesítés tapasztalatai a győri megyei könyvtárban

A győri Dr. Kovács Pál Megyei Könyvtár és Közösségi Tér a volt Galgóczi Erzsébet Városi Könyvtár és a Kisfaludy Károly Megyei Könyvtár összevonásából jött létre. Mindkét könyvtár a HunTéka integrált könyvtári rendszert használta, de mindkét adatbázis rendelkezett egyedi jellegzetességekkel és fejlesztésekkel, ezért a konverzió nem volt egyszerű. A helyzetet súlyosbította, hogy a két intézmény sok esetben eltérő gyakorlatot folytatott (feldolgozás, kölcsönzés, MOKKA-kompatibilitás megléte vagy hiánya), és mindez természetesen leképeződött az adatbázisokban. Az egyesítés egy projekt keretében zajlott le, amelyben csaknem valamennyi kolléga részt vett.

Első lépésként megvizsgáltuk az alapvető különbségeket a két rendszer működése között, majd tárgyaltunk az integrált könyvtári rendszer konverziós szakértőjével. El kellett döntenünk, hogy *melyik legyen az a primer adatbázis (és vele együtt az abhoz tartozó szoftverváltozat)*, amibe áttöltjük a másik adatait, valamint körvonalaztuk a dolog legproblémásabb vetületeit, illetve megegyeztünk, hogy *külön kezeljük az adatátviteli és a funkcionális (szoftvert érintő) feladatokat*. A Kisfaludy Károly Könyvtáré lett az elsődleges adatbázis, mivel több olyan TÁMOP-ban megvalósított fejlesztést tartalmazott (pl. névtér, amely az egyébként is problémás szerzői törzsre épül), amelyet nehéz lett volna konvertálni a másikba. Adatösszetöltési kérdésekben a legnagyobb gondot a köztaurusz verziók különbözősége, a duplumszűrés, valamint az olvasói törzsek és tranzakciók összetöltése okozta; a működés tekintetében pedig az egy- és több-lelőhelyes kölcsönzés okozta különbségeket kellett egységesíteni. Külön megoldandó feladatot jelentett, hogy az adatbázisok külső kapcsolatai (MOKKA-ODR, Duna Menti Közös Katalógus, Digitális Könyvtár, Europeana, gyarapodási RSS-ek, helyi építésű adatbázisok) továbbra is megfelelően működjenek.

Ezután következett a munka – számunkra – legnehezebb része: el kellett döntenünk, hogy a majdani egyesített adatbázisban hogyan fogunk dolgozni, és ehhez kellett igazítani a konverzió lépéseit. Az egységesítésnél a következő szempontokat vettük figyelembe: hatékonyság, MOKKA-kompatibilitás, több-lelőhelyes szolgáltatás igényei, olvasói elvárások. Először a koncepcionális kérdéseket kellett megoldani, ezután következhetek

a konkrét kérdések. Az egyeztetések csapatmunkával történtek, a projektvezető csak a végeredménnyel foglalkozott, illetve akkor avatkozott közbe, ha nem sikerült megfelelő eredményre jutni. *Négy koncepcionális kérdésben* kellett megegyezni még a konkrét munkák megkezdése előtt:

- A különböző elvek alapján épített helyismereti gyűjtemény kérdései.
- A lelőhely-gyűjtemény kategóriák eltérő alkalmazása az adatbázisban.
- A másképpen kezelt idegen nyelvű gyűjtemény kérdésköre.
- Állománymenedzselési (raktározási) problémák.

A helyismereti gyűjtemény esetében (ahol mindkét intézmény koncepciója megfelelő volt, csak éppen különböző) a volt megyei könyvtár gyakorlatát vettük át, mert az ő gyűjteményük nagyobb, így kevesebbet kellett javítani. Az idegen nyelvű gyűjtemény esetében – felhasználva mindkét könyvtár gyakorlatának pozitívumait – teljesen új koncepció született. A raktározási problémát pedig menet közben jól átgondolt és nagymértékű selejtezéssel megoldottuk. *(Lásd e számunkban a következő írást! A szerk.)* A lelőhely-gyűjtemény eltérő használatának okai:

- A volt városi könyvtárhoz több szolgáltatási hely (gyermekkönyvtár és fiókkönyvtárak) tartozott, ezért a lelőhely kategóriát a telephelyekre, a gyűjtemény kategóriát pedig a különgyűjteményekre használta.
- A megyei könyvtár a lelőhellyel csak a hangtárat különítette el a fő épülettől, és a gyűjteménnyel jelölte a dokumentumok fizikai helyét (szabad polc, alsó raktár, külső raktár, kézikönyvtár, olvasóterem stb.).

Ebből a helyzetből kellett egy olyan rendszert kialakítani, ami alapján az olvasók továbbra is megtalálják a dokumentumokat, és teljesen elkülöníthetők az adatbázisban a különböző szolgáltatási helyeken található művek, valamint lekereshetők a különgyűjtemények is. Ez persze – éppen úgy, mint a többi koncepcionális kérdés esetében – egy csomó előzetes átalakítást jelentett a két adatbázisban, hogy az összetöltés – legalábbis ebből a szempontból – majd zökkenőmentesen menjen. Ezután az egyes konkrét munkafolyamatok egységesítése következett, ami szintén egy sor előkészítő munkát igényelt mind a könyvtár dolgozóitól, mind a konverziós szakembertől.

Míndeközben küszködtünk a köztaurusz problémával is. A városi könyvtárnak egy 2005-ben betöltött, tökéletesen működő köztaurusza volt, a megyei könyvtárnak pedig egy frissített, működési problémákkal küszködő. Mivel a gondokat nem szerettük volna továbbvinni az összetöltött adatbázisba, először arra gondoltunk, hogy megtartjuk a régi, de jó köztauruszt. Aztán kiderült, hogy túl sok új, már használatba vett tárgyszó van a problémásban, amelyeket mind kézzel kellett volna javítani. Ezért úgy döntöttünk, hogy megpróbálkozunk a legújabb köztaurusz betöltésével, annak ellenére, hogy nem szerencsés az adatbázis-egyesítés bonyolult műveletét egy újdonsággal terhelni (addig még nem töltötték be a legújabb verziót egyetlen Huntéka-adatbázisba sem), mivel sok olyan új tárgyszót tartalmazott, amelyekre szükségünk volt. A problémát az okozta, hogy *az új tárgyszó-adatbázis szerkezete teljesen megváltozott, az addig egységes törzset szétszedték különböző részekre: általános tárgyszavak, formai tárgyszavak, földrajzi tárgyszavak és kronologikus tárgyszavak.* A különböző típusú tárgyszavakat tehát ezen túl nem egy mezőbe kell beemelni, hanem négy különbözőbe. Ez azért jelentett gondot, mert a földrajzi és a formai tárgyszó mezőbe eddig mi köztauruszon kívüli saját tárgyszavakat használtunk, illetve egy speciális hely-

ismereti tárgyszórendszert, amelyeket most valahogy integrálni kellett a köztauruszba, illetve megoldani – az új tárgyszótörzsben nem szereplő deskriptorok esetében – a békés egymás mellett élésüket. Mit ne mondjak, nem volt egyszerű.

A duplumszűrés problematikája volt a legfájóbb pontja az egész projektnek. Hiába találtuk meg az *ideális duplumszűrő kulcsot*, olyan választás elé kerültünk, ahol a jó választás a kisebbik rossz elfogadására korlátozódik. Azzal kellett szembesülnünk, hogy *kapcsolati rekordok esetében* (többkötetes, analitika, cikk) *vagy a kapcsolatok maradnak meg* és nem lesz duplumszűrés, *vagy lesz duplumszűrés*, de elvesznek a kapcsolatok, vagyis nem lesznek összekötve a közös és kötetrekordok, a fő- és részdokumentumok. Rengeteg kapcsolati rekordunk van, és ezekenél bizony ki kellett iktatnunk a duplumszűrést, hogy megőrizhessük a kapcsolatokat, mert azt gondoltuk, hogy könnyebb a dupla rekordokat egyesíteni, mint a szétesett kapcsolatokat újra felépíteni. Természetesen a nem kapcsolatos rekordok esetében (ezek voltak többségben) a duplumszűrés megfelelően működött.

Ezzel párhuzamosan az *olvasói törzsek egyesítésével* is foglalkoznunk kellett. Itt a *közös olvasók* okoztak gondot, akik mindkét adatbázisban szerepeltek. Úgy kellett megoldanunk a kérdést, hogy az olvasói törzsek összetöltése adatvesztés és adatkeveredés nélkül valósuljon meg. Ez kizárta a duplumszűrés alkalmazását. Az azonosnak tűnő adatok összefűlésére maradt a manuális módszer: egy táblázatba kigyűjtöttük a mindkét könyvtárba beiratkozott olvasókat mindkét vonalkódjukkal együtt. A csak egyik könyvtárba beiratkozott olvasók adatait összevonták egy adatbázisba, a táblázatba kigyűjtötteket pedig ellenőrzés után hozzáadták.

Az előkészületek után került sor a próbakonverzióra, az eredmény-adatbázis tesztelésére, a hibák javítására és a felmerülő újabb és újabb problémák megoldására. És itt jött a neheze! Az adatbázisok éles összetöltése hosszú folyamat, és nem szeretnénk volna hetekre bezárni a könyvtárat. Ezért úgy valósítottuk meg, hogy az olvasói adatokon, a kölcsönzési tranzakciókon és a statisztikákon kívül mindent nyitvatartási és kölcsönzési időben csináltunk. Számítottunk rá, hogy a rendszer sokkal lassabban fog működni, ha éles adatösszetöltés közben kölcsönzünk (ez a primer adatbázist érintette), sőt kisebb technikai fennakadásokra is fel voltunk készülve, de viszonylag zökkenőmentesen lezajlott a dolog. Ezután néhány napra bezártunk, hogy az olvasói törzseket, a kölcsönzési tranzakciókat és a statisztikákat is egyesítsük, illetve elvégezzük az első teszteket, amelyek elsősorban az olvasószolgálati munkafolyamatokra vonatkoztak, mert kinyitni csak úgy lehetett, ha ezek jól működnek. Az egyesítés után megmaradt programváltozatot a fejlesztőknek fel kellett készíteni a másik (továbbiakban nem használt) program egyedi funkciói alapján a több leülőhelyes kölcsönzésre. Ezzel gyakorlatilag a két Huntéka program funkciói is egyesültek. Javítottuk a felmerülő nem kisszámú hibát, kinyitottunk, és ekkor tudtunk foglalkozni a többi problémával. Szép lassan minden a helyére került, bár az egyesített adatbázisban maradtak hibák (az alap adatbázisokban is voltak, hiszen hibátlan adatbázis nem létezik, de természetesen a konverzió mindig generál újakat). **Végül is az egyesített rendszer működőképes és használható lett.** Természetesen még egy csomó utómunkálatra van szükség, elsősorban a dupla rekordok megszüntetésével kapcsolatban, hiszen akármilyen jó a duplumszűrő kulcs, az adatbeviteli különbségek, illetve hibák miatt sosem működhet tökéletesen.

Tanulásként levonható, hogy két adatbázis összetöltése sokkal komplexebb folyamat, mint először gondolná az ember. Bár itt csak a nagyobb volumenű munkafázisok kerültek szóba, iszonyúan sok apró részlettel kell foglalkozni, semmit nem szabad figyelmen kívül hagyni: **minden munkafolyamat, minden adatbázismező nagytít alá kell, hogy kerüljön az előkészítő szakaszban, és minden előkészítő munkát a legalaposabban kell elvégezni, ha használható végeredményt akarunk elérni.** Idő közben így is előkerülnek váratlan problémák, amelyeket meg kell oldani, de érdemes ezek számát a lehető legkisebbre redukálni. És még valamit. Egy ekkora munka nem valószínűsíthető meg néhány emberrel, **minden dolgozónak összehangoltan részt kell venni benne** a saját területén, hiszen azt ő ismeri a legjobban.

Tolnai Gáborné



Mikes Kelemen Program

A Mikes Kelemen Program munkálatai az idei, a 2015. évben is folytatódtak.

Ezek célja, hogy a diaszpórában élő magyarok megsemmisüléstől veszélyeztetett könyvtári és levéltári hagyatékait és a felajánlott magyar vonatkozású dokumentumokat rendezett módon összegyűjtve Magyarországra szállítsák, és gondoskodjanak későbbi méltó felhasználásukról. Ebben az évben a gyűjtési munkálatok kiegészültek a magyar közösségekben végzett történeti kutatásokkal is.

A program a Miniszterelnökség Nemzetpolitikai Államtitkársága irányításával, valamint az Országos Széchényi Könyvtár és a Magyar Nemzeti Levéltár szakmai felügyelete mellett zajlik. A beérkezett dokumentumokkal elsősorban az Országos Széchényi Könyvtár és a Magyar Nemzeti Levéltár gyűjteményeit egészítik ki, továbbá az anyaország és a Kárpát-medence könyvtáraiban és iskoláiban fogják elhelyezni, illetve olyan közgyűjteményekben, amelyek eddig is jelentős emigráns állományokkal bírtak.