

# Szolgáltatásfejlesztés, adatbázisépítés együttműködéssel

Burmeister Erzsébet ([erzsi@marki.lib.uni-miskolc.hu](mailto:erzsi@marki.lib.uni-miskolc.hu))

## Bevezetés – a szolgáltatás története

A MATARKA (Magyar Folyóiratok Tartalomjegyzékeinek Kereshető Adatbázisa: [www.matarka.hu](http://www.matarka.hu)) immár három éves. Fejlesztése 2002 januárjában kezdődött egy 2001-es Nemzeti Kulturális Örökség Minisztériumánál elnyert pályázati támogatás segítségével. A Miskolci Egyetem Könyvtára kapott ekkor egymillió forintot 15 magyar műszaki folyóirat 10 évfolyamának feldolgozására. A két fejlesztő három, már akkor is széles körben elterjedt szabad szoftver együttesét választotta: a MySQL relációs adatbázis-kezelőt, PHP HTML ágyazott parancsnyelvet és az APACHE Web kiszolgálót. Februárban 12 könyvtáros kollégánál már szorgalmasan vitte is be az adatokat, hogy a pályázattal el lehessen határidőre számolni. Az akkori felületre már nem is igen emlékszünk, a szerzői és cím kulcsszó indexet lehetett böngészni, az adatbázist alig használta még valaki, bár már elérhetővé vált, hisz el volt rejtve a könyvtár szerverén egy hosszú URL mögött.

A pályázat lezárásával az elkezdett munka nem szakadt meg. Megszerettük és folytattuk. Újabb NKÖM pályázat következett és most már a KATALIST könyvtárosi levelező listán megszervezett 8 tagú, különböző nagyságú könyvtárakból alakult együttműködés kapott 2,5 millió forintot a további munkához. A feldolgozott folyóiratok száma 97-re bővült. 2002 őszén hatalmas ütemben kezdett növekedni az adatbázis, ami azóta sem állt le. Természetesen a fejlesztés is folytatódott, egyre újabb lehetőségeket építettünk be, teljes egészében megújult a beviteli felület, a lekérdező felület. Több módon lehetséges a használat: az adatbázis kereshető szerző és cím kulcsszó szerint, a szerzői és kulcsszó index tanulmányozható a névváltozatok jobb áttekintése végett. És természetesen a tartalomjegyzékek folyóirat számonként böngészhetők. Mind a közvetlen kereséssel kapott találati listákból, mind a szám szerint böngészett tartalomjegyzékekből el lehet jutni a teljes szövegekhez, amennyiben azok valahol léteznek az interneten és az ugrópontok bekerültek az adatbázisba. Teljes szöveg a címek mintegy tíz százalékánál érhető el. Keresés előtt szűkítések állíthatók be: kiadási év től-ig formában, 8 szakterületre és egyenként folyóiratokra szűkítés, valamint kérhető az is, hogy csak a teljes szövegekhez elvezető ill. csak az Elektronikus Periodika Archívumban /EPA-ról később még részletesen szó lesz/ megtalálható cikkek jelenjenek meg. A szűkítések szükségesek voltak, mert a feldolgozott folyóiratok sokféle szakterületet fednek le, ezért egy-egy keresés rengeteg nem releváns találatot is eredményezhet a felhasználónak. De ha a keresés előtt bejelöli pl. a nyelvészet szakterületet, akkor csak a számára érdekes nyelvészeti szakcikkek jelennek meg a listában.

2003 elején a szolgáltatásnak saját nevet adtunk és saját domain nevet kértünk /MATARKA/, ami a használatot nagyon fellendítette. Azóta statisztikát is vezetünk a keresésekről. A letárolt keresések visszacsatolásként sok új ötletet adtak a keresési módszerek átalakítására, módosítására: alkalmazkodni kellett az interneten elterjedt általános felhasználói keresési szokásokhoz, vagyis hogy a kevésbé tapasztalt felhasználók nem ismerik a könyvtárosok számára nyilvánvalónak tűnő, Boole algebraira épülő találati halmazok egyesítésére, ill. metszetére épülő keresési stratégiákat, az átlagos felhasználó csak kereső szavakat tud egymás mellé írni. A többség számára a csonkolás is ismeretlen fogalom, arra sem gondol, hogy egy adatbázisba nem kerül be a szerzők neve elé a dr., ifj., id.

## **Az adatbázis építése**

Kézi adatbevitel böngésző segítségével: Az adatbázis biztonsága érdekében minden könyvtár felhasználói névvel és jelszóval rendelkezik a belépéshez és csak adatbeviteli, -módosítási és -törlési műveletek elvégzésére van jogosultsága. Az adatbázis rendszergazdája a Miskolci Egyetem könyvtárának dolgozója. Az adatbevitel a nyomtatott folyóirat számainak kézbevitelével űrlapok kitöltésével történik. Az adatbeviteli munkát szabályzat rögzíti.

Kézi adatbevitel egyszerű szövegszerkesztővel: A könyvtár vagy folyóirat szerkesztőség a MATARKA\_formatum szabályzat szerint veszi fel a tartalomjegyzékeket és a kész fájlokat e-mailben elküldi az Miskolci Egyetem munkatársainak, akik az erre a célra kifejlesztett program segítségével a tartalomjegyzéket az adatbázisba bekonvertálják.

Már meglevő, interneten megtalálható tartalomjegyzékek átalakítása és konvertálása: Amennyiben valamely könyvtár vagy maga a folyóirat szerkesztőség rendszeresen közzéteszi honlapján a tartalomjegyzékeket, azok onnan aránylag kevés munkával leszedhetők és egyszerű szövegszerkesztővel konvertálható formátummá átalakíthatók, akkor egy-egy szám bevitelének ideje lecsökken, hiszen nincs kézi adatbevitelre szükség, csak a szerkesztési és a konvertálási műveletsor elvégzésére.

## **Együttműködés a konzorciumon belül**

Az első két sikeres pályázat után a résztvevők száma tovább bővült, újabb tagok jelentkeztek.

Mivel azonban az utóbbi években a pályázatkírók kimondottan csak új szolgáltatások létrehozását támogatták, célzottan az adatbázis továbbfejlesztésére és továbbépítésére nem lehetett pályázatot benyújtani és sajnos más jellegű benyújtott pályázataink sem nyertek. A szolgáltatás további fenntartása érdekében szükségesnek látszott együttműködési szerződés kötése és konzorcium alakítása.

A konzorcium 2004-ben alakult az alapító könyvtárakból és csatlakozott könyvtárakból. Idézet az együttműködési szerződésből, amely július 1-én lépett életbe.

*A könyvtárak vállalják, hogy az általuk feldolgozott folyóiratok (részletes és pontos listák minden egyes könyvtár esetén az együttműködési nyilatkozatban felsorolva) tartalomjegyzékeit adatbázisba beviszik, a munkában résztvevő munkatársak az adatbázis építésével kapcsolatos tevékenységeket (elsősorban adatbevitel, de esetenként együttműködés a folyóirat szerkesztőségekkel) folyamatosan végzik. A nyilatkozatot a könyvtárak vezetői, nem önálló könyvtárak esetében a fenntartói írják alá az intézmény pecséttel ellátott céges papírján. Az együttműködési nyilatkozatok a szerződés mellékleteit képezik. A könyvtárak a feldolgozott folyóiratok listáját évente felülvizsgálják, ha változások történnek (valamely folyóirat megszűnik, adott folyóirat feldolgozását már nem tudják vállalni, újat szeretnének hozzávenni, stb.), akkor új együttműködési nyilatkozatot küldenek az Miskolci Egyetem Könyvtár Levéltár Múzeumnak, aki a konzorcium tagjait erről értesíti, és szükség esetén gondoskodik a szóbanforgó folyóiratok feldolgozásának további módjáról.*

2005 elején a konzorciumot alakító könyvtárak listája (a könyvtár neve mellett a feldolgozott folyóiratok száma).

Alapító tagok:

1. Berzsenyi Dániel Főiskola, Központi Könyvtár (5)
2. Budapest Főváros Levéltára (13)
3. Márton Áron Szakkollégium (12)

4. Miskolci Egyetem, Könyvtár, Levéltár, Múzeum (50)
5. MTA Nyelvtudományi Intézet Könyvtára (10)
6. Nyíregyházi Főiskola Központi Könyvtár és Szakirodalmi Információs Központ (15)
7. Szegedi Tudományegyetem Szegedi Élelmiszeripari Főiskolai Kar Könyvtára (17)
8. VÁTI Magyar Regionális Fejlesztési és Urbanisztikai Közhasznú Társaság, Dokumentációs Központ Könyvtár (8)

Csatlakozott tagok:

9. BME OMIKK (9)
10. DEENK Agrártudományi Könyvtár (34)
11. DEENK Társadalomtudományi Könyvtár (6)
12. MÁV Dokumentációs Központ és Könyvtár (8)
13. MTA Regionális Kutatások Központja Könyvtára (4)
14. Környezetvédelmi és Vízügyi Minisztérium Természetvédelmi Hivatal (65)
15. Kiskunsági Nemzeti Park Igazgatósága (1)
16. VMLK Környezet- és Természetvédelmi Szakkönyvtár (2)

A tagok a kapcsolatot egymással e-mailen keresztül tartják. Van egy tarka-l nevű, a Miskolci Egyetem által működtetett levelező lista. A résztvevők először 2002 szeptemberében találkoztak Budapesten a beinduló közös munka részleteinek megbeszélésére és megismerkedés végett. 2004 tavaszán ismét Budapesten tartottunk találkozót.

### **Együttműködés az Elektronikus Periodika Archívummal**

Az Országos Széchenyi Könyvtár MEK osztályán belül működik az Elektronikus Periodika Archívum /EPA/.

Az EPA célja: nyilvános hozzáférésű, magyar vagy magyar vonatkozású, tudományos és kulturális témájú elektronikus folyóiratok, magazinok, újságok és hírlevelek teljes szövegű archiválása és a magyar vonatkozású elektronikus időszaki kiadványok lelőhely és forrás-nyilvántartása.

A Miskolci Egyetem Könyvtár, Levéltár, Múzeum és az OSZK 2004 februárjában együttműködési megállapodást kötött /Februárban a MATARKA konzorcium hivatalosan még nem létezett./ Részlet az együttműködési megállapodásból:

*Az EPA és MATARKA együttműködése mind a két fél számára hasznos, mert a közös folyóiratok szerző és cím szerint kereshetővé válnak (EPA számára előnyös), a teljes szövegek megbízható ugrópontokra vezetnek és nincs szükség a konvertálásnál nagyobb munkaráfordítást igénylő kézi adatbevitelre (MATARKA számára előnyös).*

Az EPA vállalta, hogy munkatársai a MATARKA-t mindig értesítik egy újabb folyóirat archiválásáról, illetve a közös folyóiratok újabb számairól és honlapján a közös folyóiratok MATARKA-n keresztül működő keresőjét működteti, így a MATARKA-nak további nyilvánosságot biztosít. Az együttműködés kezdetén ez azt jelentette, hogy a MATARKA mindig e-mailben kapott az EPA munkatársától értesítést, ha egy új számot dolgoztak fel. Mára az értesítés automatikusan generálódik, vagyis az emberi feledékenység ki van zárva.

A MATARKA vállalta, hogy az összes EPA-ba kerülő és a gyűjtőkörébe illeszkedő folyóirat tartalomjegyzékét konvertáló módszere segítségével átveszi és az archivált teljes szövegű cikkeket belinkeli. Ez a MATARKA - szempontjából előnyös, mert az EPA garantálja a cikkek elérhetőségét. Az EPA munkatársai a közös folyóiratok listáját mindig naprakészen látják.

## Együttműködés folyóirat szerkesztőségekkel

A MATARKA böngészhető oldalait, vagyis magukat a tartalomjegyzékeket az intelligens kereső robotok /Google/ fel tudják tájni, annak ellenére, hogy azok nem statikus HTML oldalakon léteznek, hanem minden alkalommal a lekérdezéskor generálódnak. Ennek eredményeképp a cikkek a Google-ban való kereséskor is előjönnek a találatok között. Ez a tény hozzájárult ahhoz, hogy a szolgáltatás különösebben sok propaganda nélkül aránylag ismertté vált. Ezért a folyóiratok számára is kedvező, ha tartalomjegyzékeik feldolgozásra kerülnek, mert ezáltal ismertségük nő. Ez a magyarázata annak, hogy sok folyóirat szerkesztőséggel alakult már ki kapcsolatunk, akik tartalomjegyzékeiket konvertálásra alkalmas formátumban Miskolcra e-mailban elküldik. Szerkesztőségek esetében csak ezt a beviteli formát alkalmazzuk. A szerkesztőségeket vagy mi kerestük meg felkéréssel, vagy ők fordulnak hozzánk azzal, hogy szeretnének az adatbázisba bekerülni. Két folyóiratot már el is kellett utasítanunk, mert szakterületük nem illett a mieink közé. Szerkesztőségek küldik az alábbi folyóiratok tartalomjegyzékét.

1. Anyagvizsgálók lapja
2. Felnőttképzés
3. Fordítástudomány
4. Kommunikáció, Média, Gazdaság
5. Könyv, könyvtár, könyvtáros
6. Könyvtári levelező/lap
7. Korall
8. Korunk : fórum, kultúra, tudomány
9. Kútbanézők
10. Levéltári szemle
11. Levéltári közlemények
12. Modern filológiai közlemények
13. Műanyag és gumi
14. Neveléstudomány és iskolakutatás
15. Századvég
16. Új forrás

Az előbbieken alapján világosan látszik, hogy a szolgáltatás további bővítésének legígéretesebb módja további szerkesztőségek bevonása. Számukra egy-két folyóiratuk havi, kéthavi vagy negyedéves számainak tartalomjegyzékét elkészíteni és nekünk elküldeni nem igazán időigényes feladat. Különösen akkor nem, ha a tartalomjegyzék valamilyen számítógépes formátumban amúgy is létezik, és ez ma már szinte természetes. Esetünkben igaz tehát a mondás: sok kicsi sokra megy. Sok ember kevés munkájából EGYÜTTMŰKÖDÉS-sel komoly és jól használható adatbázis épül.

## Tervezett együttműködés az EPA-val és LibInfo-val

2004-ben az EPA, a LibInfo (az OSZK szolgáltatása) és a MATARKA tervezte, hogy közös GVOP pályázatot ad be, a pályázat el is készült, de a túl magas kötelező önrész miatt az OSZK nem vállalta a benyújtást. A megvalósítandó merész és újszerű tervek egyike az lett volna, hogy a MATARKA-ban a teljes szöveggel el nem érhető cikkeknek a felhasználó a kívánt cikk fénymásolatát SMS-en keresztül történő fizetéssel megrendelhetné volna a folyóiratért felelős könyvtártól, a beszkenelt cikket e-mail címére kapta volna meg, illetve adott honlapról letölthette volna. Talán lesz még egyszer ebből is valami....

## Néhány statisztikai adat (2005 február 6.)

A statisztikák a honlapról is elérhetők mindig.

Adatbázis nagysága

Folyóiratok száma: 274

Cikkek száma: 182893

Szerzős cikkek száma: 138111

Szerzők száma: 47416

Nem üres címsorok száma: 178955

URL-ek száma: 18265

Látogatók száma 2004-ben:

Hónap	Keresés	Böngészés	Tallózás	Összesen
Január	5358	3137	690	9185
Február	8402	3875	1195	13472
Március	10206	3069	1404	14679
Április	9546	2077	1139	12762
Május	6205	2381	1144	9730
Június	4133	2407	557	7097
Július	2798	1949	461	5208
Augusztus	2648	2644	542	5834
Szeptember	7495	4023	1071	12589
Október	12810	5726	1609	20145
November	15213	4303	1706	21222
December	10342	4093	1072	15507
2004 összesen	94929	39637	12556	147122