

Hogyan tovább MATARKA?

Burmeister Erzsébet

informatikai osztályvezető, Miskolci Egyetem, Könyvtár, Levéltár, Múzeum

(A Könyvtárosok Vándorgyűlésén Miskolcon, 2004 július 30-án elhangzott előadás)

A MATARKA /MAgyar Folyóiratok TARtalomjegyzékének Kereshető Adatbázisa/ internetes szolgáltatás, a www.matarka.hu internetes oldalról érhető el. Az előadás keretében szó lesz a MATARKA szolgáltatásról, a történetéről, bemutatom a résztvevő könyvtárakat, statisztikákat ismertetek, illetve választ keresek arra a kérdésre, hogyan lehet az adatbázist és a hozzá kapcsolódó szolgáltatást bővíteni és tovább fejleszteni.

1. Mi az a MATARKA?

Részlet a súgóból a www.matarka.hu alatt: *"A MATARKA magyar kiadású szakfolyóiratok tartalomjegyzékét dolgozza fel könyvtári együttműködés keretében a Miskolci Egyetem Könyvtár, Levéltár, Múzeum vezetésével. A felvett adatok adatbázisba kerülnek és több szempont szerint visszakereshetők ill. az egyes számok külön-külön böngészhetők.*

Ha a /megtalált, böngészett/ cikk(ek) az interneten teljes szöveggel elérhető(k), akkor a teljes szöveghez ugrópont(ok) vezet(nek). Ha a cikk(ek) az interneten teljes szövegű adatbázisban elérhető(k), de a cikk(ek)hez külön ugrópont nem adható(k) meg, akkor a folyóirat címéről vezet ugrópont a teljes szövegű adatbázis URL-jéhez.

A feldolgozás egyrészt kézi adatbevitellel történik, másrészt digitális formában már rendelkezésre álló tartalomjegyzékek konvertálásával."

2. A MATARKA története dióhéjban

2001-ben a Miskolci Egyetem Könyvtár, Levéltár, Múzeum a Nemzeti Kulturális Örökség Minisztériumához (NKÖM) benyújtott pályázatával (*Magyar kiadású műszaki és természettudományi folyóiratok tartalomjegyzékének internetes szolgáltatása egy kereshető és böngészhető rendszerben*) 1 millió forint támogatást nyert, melynek segítségével 2002 januárjában elindulhatott a munka. Kiss Andrea, aki szintén a könyvtár informatikai osztályának dolgozója és jelen előadás tartója elkezdte a szolgáltatás koncepciójának kidolgozását és egy időben magát a megvalósítást is. Megterveztük az adatbázist a MySQL szabadon használható adatbázis kezelő szoftver igénybevitelével, a PHP, HTML-be ágyazható programozási nyelv segítségével pedig elkészítettük az adatbázis feltöltésére szolgáló beviteli űrlapokat, majd a lekérdező felületet. Az adatbázis kezelésére szolgáló programokat is megírtuk, amelyek segítségével 12 miskolci könyvtáros már 2002 februárjában elkezdte az adatbeviteli munkát, 15 műszaki folyóirat 10 évre visszamenő tartalomjegyzékének feldolgozását.

2002-ben a könyvtár újabb pályázatot nyújtott be a NKÖM-höz (*Tartalomjegyzék szolgáltatás további bővítése, országos tartalomjegyzék szolgáltatás létrehozása*) most már 7 másik könyvtárral együtt, amelyek a Katalist levelező listán közzétett felhívásra jelentkeztek. A minisztériumtól most két és fél millió forint támogatást kaptunk. Az együttes munka 2002 októberében egy Budapesten megtartott személyes találkozóval kezdődött, ahol megismertük egymást, megbeszéltük az adatbevitelt, a gyakran előforduló nehézségeket, egyeztettük a feldolgozási szabályokat.

A továbbiakban egyrészt sor került a szoftver finomítására, ami azóta is folyamatos munkát jelent, hisz állandóan újabb és újabb szempontokat igyekszünk a szolgáltatás fejlesztésénél figyelembe venni, ill. újabb és újabb igényeket kívánunk kielégíteni, másrészt több könyvtáros kezdte el az adatbevitelt, így az adatbázis nőtt, a feldolgozott folyóiratok száma bővült. Ez a fejlődés azóta is töretlen annak ellenére, hogy a kétszeri sikeres pályázás óta újabb támogatást máshonnan már nem kaptunk, pályázatainkat visszautasították.

Fontos dátumok felsorolásszerűen a MATARKA történetében létrejöttétől, 2002 januárjától kezdve, beleértve a megjelent publikációkat és elhangzott előadásokat:

- | | |
|--------------|---|
| 2002 január | beviteli felület fejlesztése |
| február | 12 könyvtáros elkezdte az adatbevitelt |
| nyár | új pályázat beadása (7 résztvevő + Miskolc) |
| augusztus | Debrecenben előadás az Informatika a felsőoktatásban konferencián (<i>Tartalomjegyzék szolgáltatás MySQL és PHP segítségével</i>) |
| október | 8 könyvtár I. találkozója a Márton Áron Szakkollégiumban Budapesten |
| ősz | szűkítések beépítése a keresésben, közös munka elkezdése, konvertáló program készítése számítógépes formában már meglévő tartalomjegyzékek közvetlen adatbázisba vitelére |
| | DAT konferencián előadás (<i>Tartalomjegyzék szolgáltatás a Miskolci Egyetem könyvtárában</i>) |
| 2003 január | MATARKA domain név regisztrálása |
| február | keresési statisztika, találati halmaz több oldalas megjelenítése |
| | Tudományos Műszaki Tájékoztatásban cikk (<i>Tartalomjegyzék szolgáltatás nyolc könyvtár együttműködésével</i>) |
| tavaszi | NETWORKSHOP konferencián előadás (<i>Nyolc könyvtár közös adatbázis építésének tapasztalatai</i>) |
| | Könyvfesztiválon előadás (<i>MATARKA - Magyar folyóiratok tartalomjegyzékeinek kereshető adatbázisa</i>) |
| | kereső felület megújulása: keresés, tallózás, böngészés; számláló bevezetése a honlapon, találati halmaz rendezése |
| nyár | újabb könyvtárak csatlakoznak |
| ősz | új beviteli felület kialakítása |
| december | Könyv, könyvtár, könyvtáros cikk (<i>MATARKA - magyar folyóiratok tartalomjegyzékeinek kereshető adatbázisa</i>) |
| 2004 február | az OSZK-beli EPA-val (Elektronikus Periodika Archivummal) |

	együttműködési megállapodás megkötése
március	cikkbibliográfiák USMARC-HUNMARC letöltési lehetősége a böngésző felületről
április	14 résztvevő könyvtár II. találkozója a Márton Áron Szakkollégiumban Budapesten
július	14+1 résztvevő könyvtár konzorciumot alapít

3. A MATARKA konzorcium résztvevő könyvtárai

Alapító tagok (8)

1. Berzsényi Dániel Főiskola Központi Könyvtára
2. Budapest Főváros Levéltára könyvtára
3. Márton Áron Szakkollégium könyvtára
4. Miskolci Egyetem Könyvtár, Levéltár, Múzeum
5. MTA Nyelvtudományi Intézet könyvtára
6. Nyíregyházi Főiskola Központi Könyvtár
7. Szegedi Élelmiszeripari Főiskolai Kar Könyvtára
8. VÁTI Magyar Regionális Fejlesztési és Urbanisztikai Közhasznú Társaság könyvtára

Csatlakozott tagok (7)

1. BME OMIKK
2. DEENK, Társadalomtudományi Könyvtár
3. DEENK, Agrártudományi Könyvtár
4. Kiskunsági Nemzeti Park Igazgatósága könyvtára
5. MÁV Dokumentációs Központ és Könyvtár
6. MTA Regionális Kutatások Központja
7. Természetvédelmi Hivatal könyvtára

4. Az adatbázis nagyságát jellemző statisztikai adatok (2004. július 22-i adatok)

Folyóiratok száma:	233
Füzetek száma:	9713
Szerzők száma:	41345
Cikkek száma:	153181
Nem üres címsorok száma:	149491
Szerzős cikkek száma:	114628
Ugrópontok száma teljes szövegekre:	11053

Magyarázat a fenti felsoroláshoz: A cikkek száma és a nem üres címsorok száma közötti különbség azt jelenti, hogy a "cikkekhez" odasoroljuk a tényleges címmel nem rendelkező rovatokat is. A MATARKA-ban a cikk tulajdonképpen a tartalomjegyzék egy sora, egy rekordja. Ha egy rovaton belül több cikk is szerepel, akkor a rovatcím az első cikknél kerül csak bevitelre, vagyis a rovat nem számolódik be a cikkek számába, nem alkot külön rekordot. De ha a rovaton belül nincs önálló cím, akkor a rovat maga a tartalomjegyzék egy egysége (egy önálló rekordja).

5. Az adatbázis használatának statisztikai adatai

Az adatbázis kereshető szerzők és a cíkcímekben szereplő kulcsszavak szerint. Ez a **keresés**. Ezen kívül mód van **tallózásra**, vagyis a szerző ill. cím kulcsszó indexekbe való belépésre. A **böngészés** az egyes folyóiratszámok tartalomjegyzékének megtekintését jelenti.

A honlapon a statisztikák menüpont alól az adatbázis használati statisztikái érhetők el 2004. január 1-től napi, havi és éves bontásban, illetve az éppen utolsó 150 keresés is kilistázható. Ez a lista felsorolja a számláló állását, az adatbázis használatának dátumát pontos időponttal és magukat a keresett, tallózott szavakat és neveket valamint a böngészett folyóiratokat.

A statisztikai adatokból fontos következtetések vonhatók le a használatra, a leginkább keresett témákra vonatkozóan, nyomon követhetők a keresési szokások (keresési hibák), amelyek alapján a további bővítés és fejlesztés tervezhető. A statisztikai adatok vizsgálata eredményezte 2003-ban a kereső kifejezés több szóra való felbonthatóságának bevezetését (Google minta !!!).

97582 --- 04Sep07-21:59:04 --- Keresés: 1-szerzo: kosinski 2-szerzo: a festett madár
97582 --- 04Sep07-21:59:39 --- Keresés: 1-szerzo: hemingway 2-reszlet: malcolm
97583 --- 04Sep07-22:04:52 --- Böngészés: Gazdaság és jog
97585 --- 04Sep07-22:21:06 --- Keresés: 1-cim: rekettye 2-szerzo:
97586 --- 04Sep07-22:22:36 --- Böngészés: Marketing & menedzsment
97586 --- 04Sep07-22:25:39 --- Böngészés: Marketing & menedzsment
97586 --- 04Sep07-22:27:07 --- Böngészés: Marketing & menedzsment
97586 --- 04Sep07-22:27:21 --- Böngészés: Egészségügyi gazdasági szemle
97587 --- 04Sep07-22:34:50 --- Böngészés: Halászat
97588 --- 04Sep07-23:26:45 --- Böngészés: Gazdaság és jog
97588 --- 04Sep07-23:27:23 --- Tallózás: szerzo: tattay
97588 --- 04Sep07-23:29:03 --- Tallózás: szerzo: tattay

1. ábra: 'Az utolsó keresések kilistázva' gomb alól ilyen lista érhető el.

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	2118	1506	378	4002
2004 összesen	45741	18405	6473	70619

2. ábra: A 2004 év használati statisztikája havi bontásban, a keresést, böngészést és tallózást megkülönböztetve.

Az éves statisztikából világosan látható, hogy a MATARKA használata az iskolaévhez kötődik. Azokban a hónapokban, amikor a felsőoktatásban vagy akár a középiskolai oktatásban a legnagyobb a hajrá és a legnagyobb szükség van az irodalomkeresésre, akkor nő meg ugrásszerűen a használat.

Sajnálatos annak megállapítása, hogy alacsony a tallózás aránya, pedig a tallózható indexekből látható például legjobban a szerzők neve sorrendben, vagy egy szó különböző ragozott alakjai. Gyakran előfordul, hogy a tartalomjegyzék csak a szerző keresztnévének kezdőbetűjét adja, ezért az adatbázisban is sokszor előfordul, hogy szerzők teljes keresztnévvel is és kezdőbetűvel megadott keresztnévvel is szerepelnek. Az egységesítés bár ajánlott, de nem előírás a MATARKA feldolgozó könyvtárosai számára.

[Tisza M.](#) (4 cikk)

[Tisza Miklós](#) (17 cikk)

[Tisza Miklósné](#) (1 cikk)

3. ábra: Részlet a szerzőnevek tallózásából

6. Hogyan tovább ?

A MATARKA bővítése jelenti az adatbázis, illetve a szolgáltatás bővítését:

- szakterületek bővítése

Adatbázis bővítése - folyóiratok számának növelése

- retrospektív feldolgozás

- összetett (Boole) keresés

Szolgáltatások bővítése - dokumentumküldés (cikkmásolatok közvetlenül)

Egy életképes szolgáltatás fejlesztése sosem állhat le, hisz a technikai fejlődéssel a lehetőségek módosulnak, és változnak a felhasználói igények is. A MATARKA esetében nagy az igény (még meg nem valósított) a teljes szövegek biztosítására. Ahol ugrópont vezet a teljes szövegre (jelenleg 40 folyóiratnál teljesen vagy részlegesen érhető el teljes szöveg), ott ez a kérdés megoldott, de nem minden folyóirat szerkesztőség olyan nagy vonalú - ami természetesen érthető -, hogy teljes szövegeit kitegye az internetre. Ezért a jövő legnagyobb feladata számunkra a dokumentumküldő szolgáltatás kifejlesztése.

Az összetett keresés ('and', 'or' és 'not' operátorok használata a keresőszavak között) bevezetés előtt áll, az előmunkálatok már megtörténtek.

A feldolgozott folyóiratok száma szakterületenkénti bontásban:

agrár

45

általános

5

gazdasági	21
jogi	6
műszaki	31
nyelvtudomány	11
társadalomtudományok	55
természettudományok	62

Jónak mondható az agrár, a nyelvtudomány, az élővilág, természetvédelem (természettudományok alatt) szakterületek lefedettsége. A jogi, gazdasági, műszaki, társadalomtudományok szakterületeket sok folyóirattal lehetne és kellene bővíteni.

A szakterületeket bővíteni egyelőre nem fogjuk. Ezért kimaradnak olyan nagy szakterületek, mint művészetek, sport.

A retrospektív feldolgozás folytatását a résztvevő könyvtárak maguk döntenek el. Ha van rá kapacitásuk, akkor csinálják. Természetesen itt az is megfontolandó, hogy adott folyóirat esetében mennyire érdemes időben visszamenni. Különösen műszaki és gazdasági folyóiratoknál jelentős az elavulás, elévülés.

7. Csatlakozás a MATARKA-hoz

A MATARKA konzorciumhoz újabb könyvtárak csatlakozását várjuk. Ezért a honlapról a dokumentumok menüpont alól a MATARKA dokumentumok letölthetőek:

1. MATARKA konzorciumi szerződés
2. MATARKA adatbeviteli formátum
3. Feldolgozási szabályok és használati útmutató a MATARKA adatbeviteli felületéhez

Csatlakozzon!