

## A magyar szakirodalom tartalmi reprezentációjának jelenlegi állapota

*A tanulmány – illeszkedve „A tudás és az információfeldolgozás” c. cikkben elemzett vizsgálathoz – a magyar szakirodalmi adatbázisok tartalmi feltáró rendszerének elemzését végzi el. Tizenhat adatbázist vizsgál meg a tartalmi feltáró ismérvek száma, forrása, jelölése, tartalma és a szövegösszefüggéshez való viszonya szempontjából. Nemzetközi összehasonlításban megállapítható, hogy a magyarországi szakirodalmi adatbázisok jelenlegi feltáró rendszere nem elegendő a részletesebb szövegrepresentáláshoz. A tudományos szakirodalom megérdemelné, hogy igényesebb szolgáltatásokat nyújtó adatbázisok álljanak a publikációk mögött, amelyek talán egy igényesebb információkereső közönséget is kitermelnének. A színvonalas információs szolgáltatások megtermékenyítőleg hatnának magára a szakirodalomra is, hiszen lehetőség lenne a minőségi válogatásra. Jelenleg ezt a válogatást az adatbázisok nem támogatják. (A tanulmány a szerző PhD-disszertációjának része).*

Magyarországon még nem igazán honosodott meg a szakirodalmi adatbázisok rendszere, nincsenek független adatbázis-előállító és -szolgáltató intézmények. Egyedül talán az Arcanum adatbázis-előállító cég kivétel, amely sok nagyszerű adatbázisa sorában szakirodalmi adatbázisokat is szolgáltat (pl. a szociológiai adatbázist és a PRESSDOK-ot), mégpedig kiváló színvonalon. Az Arcanum adatbázisai a kellően szegmentált adat-szerkezetet és az ehhez szükséges korszerű keresőeszközöket egyaránt biztosítják. Általában azonban a könyvtárak vállalják a feladatot, és többnyire saját katalógusaik keretein belül szolgálják ki az információs igényeket. A vizsgálatot a tudományos és szakkönyvtárak feltáró munkájának elemzésével célszerű végrehajtani, hiszen ezek azok az intézmények, amelyeknek feladata a megfelelő mélységű tartalmi feltárás.

Mivel kicsi a piac – a nemzetközi adatbázisok piacához viszonyítva –, a kifinomultabb adatbázisok előállításához nincsenek megfelelő szoftverek sem. Ennek természetesen indoka, hogy nem érdemes nagyon költséges szolgáltatásokat nyújtani, amelyeket nem tud a kereslet eltartani. A szöveghez igazodó szurrogátumszerkezetet a könyvtári katalógusokat (OPAC-okat) szolgáltató integrált könyvtári szoftverek nem igazán támogatják. Lehetőség van kulcsszavas feltárássra (vagyis a cím szignifikáns szavai alapján történő keresésre), hierarchikus és mellérendelő osztályozások,

tárgyszavas feltárás alkalmazására, szintaxisra azonban az esetek többségében nincs mód. Ugyancsak kevés referátum készül a magyar tudományos szövegekről. Azt kell mondani, hogy a felhasználóknak nem sok lehetőségük van arra, hogy információs igényeikre gyors és célzott választ kapjanak, inkább arra kell felkészülniük, hogy nagyobb mennyiségű szövegből kell kiválogatniuk a relevánsakat.

A nemzeti szakirodalom feltárásának teljességét a nemzeti szakbibliográfiák rendszere valósítja meg. Egy ország, illetve nemzet publikációs vagyont a nemzeti bibliográfiák veszik számba, ezeknek azonban ténylegesen csak a számbavétel és a regisztráció a feladatuk. Nem hivatottak tehát arra, hogy tartalmi alapú tájékoztató szolgáltatásokat, információszerkezetet építsünk rájuk. Ezt a feladatot az egyes tudományterületek információiért felelős országos vagy regionális intézmények valósítják meg. Vizsgálatom arra irányul, hogy az országos szakirodalmi információellátásért felelős intézmények elektronikusan hozzáférhető szolgáltatásai milyen szinten valósítják meg a 21. század elején Magyarországon a szövegek információ-tartalmának korszerű reprezentációját.

A kulturális javak védelméről és a muzeális intézményekről, a nyilvános könyvtári ellátásról és a közművelődésről szóló 1997. évi CXL. törvényben felsorolt országos szakkönyvtárak szolgálta-

tásai között a következő szakirodalmi adatbázisokat találjuk. A sort kiegészítettem még néhány más intézmény érdekes, színvonalas adatbázisával.

#### **Budapesti Műszaki Egyetem Országos Műszaki Információs Központ és Könyvtár (BME OMIKK)**

A műszaki és élettelen természettudományok országos szakkönyvtára. 2001-ig előállította a műszaki és természettudományi cikkeket feldolgozó adatbázisát (CIKK), amely kulcsszavakkal, kötött és szabad tárgyszavakkal, valamint hárombetűs kódokból álló osztályozási jelzetekkel tárja fel a szövegek tartalmát. Folytatása a Magyar Műszaki és Gazdasági Repertórium adatbázisa (MMR), amelyben már csak magyar cikkek találhatóak, ugyanakkor azonban a konferenciakötetek tanulmányait is feltárja. Tartalmi feltáró eszközei ugyanazok.

#### **Országos Egészségügyi Információs Központ és Könyvtár (MEDINFO)\***

Az egészségügy, az orvostudomány, a gyógyszerészet és az ápolástudomány országos szakkönyvtára, a Magyar Orvosi Bibliográfiát és adatbázist (MOB) állítja elő. Az adatbázis könyveket, folyóiratcikkeket és más típusú dokumentumokat egyaránt feldolgoz. A szövegek tartalmi feltárására kulcsszavakat, szabályozott tárgyszavakat és speciális osztályozási jelzeteket alkalmaz, de ezek nem mindig a tartalomra utalnak. A folyóiratcikk egy részéhez tartalmi kivonat is tartozik, ezek azonban nem a szorosán vett szakcikkek, inkább ismeretterjesztő jellegűek. Tárgyszórendszere a National Library of Medicine MESH rendszerének magyar nyelvű adaptációja, a szókészlethez tartozó indexelési elveket azonban nem vették át.

#### **Központi Statisztikai Hivatal Könyvtár és Dokumentációs Szolgálat**

A statisztika országos szakkönyvtára. A TEXTPAC 1977 óta folyamatosan épülő számítógépes szakirodalmi adatbázis, hazai és külföldi tudományos folyóiratok cikkeinek visszakeresési lehetőségével. Az adatbázis lefedi a hazai és a nemzetközi társadalmi-gazdasági élet szinte minden területét. Keresni lehet szerző, kulcsszó, földrajzi kód, nyelv és szabad tárgyszavak szerint.

#### **Országgyűlési Könyvtár**

A jogtudomány, közigazgatás, a politika és a 20. századi egyetemes történelmi tudományok országos szakkönyvtára. Több adatbázisban is feltárja a jogtudomány és a politika szakirodalmát:

- *Magyar jogi szakirodalmi adatbázis, 1867–1944.* Tartalma: a magyar jog válogatott irodalma. Könyvek, folyóiratcikk, tanulmánykötetek tanulmányai, 1867-től 1944-ig, kiegészítve néhány korábbi tétellel. A tételek a szerző, szerkesztő, összeállító stb., a cím szavai, a közzétevő intézmény, a folyóirat címe, raktári jelzete, a könyv kiadási helye és éve, valamint nyelv szerint kereshetők. Válogatott tárgyszavak szerinti keresésre is lehetőség van.
- *Magyar jogi szakirodalmi adatbázis, 1945–1979:* A magyar jog válogatott irodalma. Könyvek, folyóiratcikk, tanulmánykötetek tanulmányai, 1945-től 1979-ig. A tételek a szerző, szerkesztő, összeállító stb., a cím szavai, a közzétevő intézmény, a folyóirat címe, raktári jelzete, a könyv kiadási helye és éve, valamint nyelv szerint kereshetők. A műfaj és téma szerinti keresésre még nincs lehetőség.
- *Magyar jogi szakirodalmi adatbázis, 1980–1989:* A magyar jog válogatott irodalma. Könyvek, folyóiratcikk, tanulmánykötetek tanulmányai, 1980-tól 1989-ig. (A magyar állam- és jogtudományi irodalom bibliográfiája c. kiadvány 1980–1989. évi köteteinek feldolgozása.) A tételek a szerző, szerkesztő, összeállító stb., a cím szavai, a közzétevő intézmény, a folyóirat címe, raktári jelzete, a könyv kiadási helye és éve, műfaj, téma (tárgyszavak) és nyelv szerint kereshetők.
- *Magyar jogi szakirodalmi adatbázis, 1990-től:* A magyar jog és határterületeinek válogatott irodalma. Az adatbázis az 1990. január 1. óta Magyarországon megjelent és a könyvtárba beérkezett magyar és idegen nyelvű jogi könyveket, tanulmányköteteket, önállóan megjelent tanulmányokat, a jogi szakfolyóiratok teljes anyagát, és más társadalomtudományi folyóiratok jogi tárgyú cikkeit tárja fel. A jogi könyvek között az önálló kötetben megjelenő jogszabályokat és jogszabálygyűjteményeket is feldolgozzák. A tanulmánykötetek feltárása általában analitikusan történik. A napi- és hetilapokban megjelenő jogi tárgyú cikkeket az adatbázis nem tartalmazza. Ilyen kérdésekre a könyvtár PRESSDOK adatbázisa ad választ. A tételek a szerző, szerkesztő, összeállító stb., a cím szavai (kulcsszavak), a közzétevő intézmény, a folyóirat címe, raktári jelzete, a könyv kiadási helye és éve, műfaj, téma (tárgyszavak) és nyelv szerint kereshetők. Az 1996. évi tételektől kezdve a hivatkozott joganyagra is lehet keresni.

\*Időközben az intézmény átalakult. Részletesebben lapunk 355–358. oldalán. (A szerk.)

- **Külföldi jogi szakirodalmi adatbázis:** A külföldi jog válogatott irodalma. Az adatbázis a könyvtár állományából kiválasztott több mint 300 külföldi jogi szakfolyóirat, valamint tanulmánykötet válogatott cikkeit és tanulmányait tárja fel, főleg 1980 óta (kiegészítve néhány jogszabály és nemzetközi egyezmény feldolgozásával). A tételek 66 százaléka csak a bibliográfiai adatokat tartalmazza. Az anyag 34 százaléka viszont ezeken kívül az eredeti cím magyar fordítását, valamint a visszakeresést megkönnyítő tárgyszavakat és kódokat is tartalmazza. Egyes tételeknél a cikk rövid ismertetése is megtalálható. A tételek a szerző(k), a cím szavai, a folyóirat címe, a kiadás éve, nyelv, valamint a tételek kisebb hányadánál téma (tárgyszavak és kódok) szerint is kereshetők.
- **Hazai sajtófigyelés (PRESSDOK):** A bibliográfiai adatbázis a hazai politikai, gazdasági és részben jogi sajtóban (mintegy 150 napi- és hetilapban, folyóiratban) megjelent cikkek szelektív feltárással épül. Segítségével naprakészen nyomon követhetjük a belpolitikai és gazdasági folyamatokat, vezető politikai, gazdasági, kulturális személyiségek nyilatkozatait, valamint a pártok, a politikai, gazdasági és érdekvédelmi szervezetek állásfoglalásait, értékeléseit. Jelenleg közel 700 000 tételből áll, havi átlagban 7-8000 cikk adatával bővül. Az adatbázis kezdő éve 1989. Keresni lehet kulcsszavak, tárgyszavak, témakódok alapján, és az adatbázis a cikkekről rövid referátumot is ad.
- **Külföldi sajtófigyelés (HUNDOK):** A bibliográfiai adatbázis forrása az Országgyűlési Könyvtár állományában megtalálható kb. 50 mértékadó külföldi – nyugat- és kelet-közép-európai, amerikai, kanadai és ausztráliai – napi- és hetilap. A kiválasztott cikkekről rövid tartalmi leírás készül. A feltárt cikkek bemutatják, mit tart fontosnak a világ Magyarországról és a kelet-közép-európai régió politikai és gazdasági eseményei közül. Az adatbázis jelenleg több mint 85 000 tételt tartalmaz, havonta átlagosan 800–1000 cikk adataival bővül, 1991-től építik.
- **Világpolitikai dokumentáció és információ (VPI):** Mintegy 100 külföldi politikai és gazdasági szakfolyóiratban megjelent cikk, valamint gyűjteményes mű, tanulmánykötet részletes (annotált) feltárással készült szakirodalmi adatbázis. Jelenleg közel 27 000 címet tartalmaz. Havonta átlagosan 200 tétellel bővül.

### **Országos Mezőgazdasági Könyvtár és Dokumentációs Központ**

A mezőgazdaság, az agrártudományok, az élelmiszeripar, az erdészet és a faipar országos szak-

könyvtára. Előállítja a Magyar Mezőgazdasági Bibliográfiát (MAB), amely folyóiratcikkek, könyvek, illetve könyvrészletek és konferenciaanyagok leírásait tartalmazza, és adatbázisként is hozzáférhető. A témára utaló adatelemek: kulcsszó, tárgyszó, ETO-jelzet, tárgyi melléktétel, tárgykód. A tárgykód kivételével mindegyik kereshető.

### **Országos Idegennyelvű Könyvtár**

Az irodalom, az idegen nyelvek és a nemzetiségi szakirodalom országos szakkönyvtára.

- **Nemzetiségi adatbázis:** A nemzeti és etnikai kisebbségekre vonatkozó irodalom válogatott bibliográfiája. A számítógépes adatbázis CD-ROM-on 1992-től 1997-ig tartalmazza a nemzetiségi-etnikai témakörben megjelent újságcikkeket, értekezéseket és az elektronikus média közléseit. Az adatbázis jelenleg 34 000 sok szempontúan feltárt és lekérdezhető tételt tartalmaz. Segítségével, mint forrástájékoztató eszközzel, az országos információszolgáltatás kisebbségi kérdésekkel foglalkozó részét kívánják ellátni. A cikkek tartalmát kulcsszavakkal, ETO-ból képzett tárgyszavakkal, osztályozási jelzetekkel és földrajzi nevekkel tárják fel, és keresni lehet a hivatkozott tulajdonvevekre is.

### **Országos Pedagógiai Könyvtár és Múzeum**

A neveléstudomány és az oktatáspolitikai országos szakkönyvtára. Nemzeti szakkönyvtári feladatként gyűjti a magyar közoktatás tankönyveit. A neveléstudomány szakirodalmi adatbázisát állítja elő.

- A **PAD (Pedagógiai Adatbázis)** 1989-től tartalmazza a Magyarországon megjelent pedagógiai szakirodalmat (könyvek, tanulmányok, folyóiratcikkek), ezenkívül egy reprezentatív válogatást a Magyarországon hozzáférhető jelentősebb idegen nyelvű folyóiratcikkekről, illetve az Országos Pedagógiai Könyvtár állományában lévő idegen nyelvű könyvekről (1975-ig visszamenőleg). Az információk elérését jól strukturált adatszerkezet és változatos keresési lehetőségek segítik. A tartalmi feltárás több szinten történik. Valamennyi tartalmi információt hordozó adatmező (pl. cím, alcím, sorozat) szavanként kereshető, a keresőszoftver lehetőségeit figyelembe véve különféle logikai kombinációkkal (és, vagy, de nem, közelségi operátorok). A mélyebb tartalmi feltárást egy átfogó szakcsoportrendszer és az Országos Pedagógiai Könyvtárban kifejlesztett tárgyszórendszer tárgyszavai adják. A szakcsoportok egyfajta osztályozási rendként a pedagógia nagyobb témaköreit ölelik fel, használatuk a keresésben egyrészt egy téma összefoglaló irodalmának megtalálását, másrészt a tárgyszavas keresés további fi-

nomítását teszik lehetővé. Az adatbázisban szereplő tételek tartalmáról a tárgyszavak és az idegen nyelvű cikkek esetében a tartalmat összefoglaló magyar nyelvű referátumok adnak képet. A tárgyszavazás egységes elvek alapján, szabályozott szótár segítségével történik, a tárgyszavak egymás utáni sorrendje az esetek többségében mondatszerűen tükrözi a közöttük fennálló szintaktikai kapcsolatokat is. Keresni lehet földrajzi nevek, iskolatípusok, iskolák neve stb. alapján, a hatékonyság azonban megkívánja, hogy a keresésben alkalmazott elemeket a keresőrendszer kínálta lehetőségek szerint kombináljuk. Így például, ha az általános iskoláról mint iskolatípusról keresünk irodalmat, célszerű az általános iskola tárgyszót a 222-es (iskolatípusok) szakcsoportszámmal kombinálni, ha azonban az általános iskola történetéről szeretnénk olvasni, akkor a 64-es (neveléstörténet) szakcsoportszámot kell a keresést szűkítő szempontként alkalmaznunk. A referátumok összefoglalják az idegen nyelvű művek tartalmát, így könnyebben eldönthető, valóban a minket érdeklő témáról van-e szó, érdemes-e kézbe venni az eredeti művet.

#### **Fővárosi Szabó Ervin Könyvtár**

Közkönyvtár és egyben a szociológia országos szakkönyvtára 1968 óta. A szociológia és a vele határos szakterületek magyar és külföldi irodalmának részletező feltárásán alapuló adatbázisokat épít.

- A Fővárosi Szabó Ervin Könyvtár szociológiai dokumentációs osztálya 1972 óta évente jelentette meg nyomtatásban *Szociológiai Információ* című bibliográfiáját, 1987 óta folyamatosan építi adatbázisát, mely több mint 100 000 bibliográfiai tételből áll, folyóiratcikkek, cikksorozatok, monográfiák, tanulmánykötetek és többkötetes könyvek leírását tartalmazza. Az adatbázis évente hozzávetőleg 200-féle folyóirat és több mint 1000 könyv analitikus feltárása alapján készül. A tételek minden lényeges adata külön-külön és tetszőleges kombinációkban visszakereshető. Lehetőség van az egy vagy több folyóiratban megjelent viták figyelemmel kísérésére, és az egyes könyvekről írott könyvismertetések áttekintésére. Visszakereshetők a részletesen feldolgozott tanulmánykötetek egyes tanulmányai, és a cikksorozatokba, tematikus számokba tartozó cikkek adatai. A tematikus keresést átfogó témakörök, a szociológiát és határterületeit lefedő átfogóbb és a részletezőbb fogalmakat leíró, javarészt tezauszba rendezett tárgyszavak, valamint a címetek helyenként kiegészítő rövid anotációk segítik. Az adatbázis nemcsak a ma-

gyarországi, hanem a külföldön megjelenő magyar vonatkozású irodalmat is tartalmazza.

#### **A Magyar Filmintézet Könyvtára**

A filmtudomány és filmművészet országos szakkönyvtára. A szakterület teljes hungarika szakirodalmát, a könyveket, folyóiratokat, forgatókönyveket és apró nyomtatványokat gyűjti.

- *A magyar nyelvű filmszakirodalom adatbázisa:* A magyar nyelvű filmszakirodalom közel kilenc évtizedes története folyamán olyan ismeretanyagot halmozott fel, amely a filmszakma, a filmtörténet, a filmkultúra megismerésének nélkülözhetetlen forrása. Sajátságos jellegére utal, hogy seregnyi adat, információ más forrásból nem szerezhető meg. Ez indokolja e forrásanyag minden részletre kiterjedő tárgyszavak rendszerére épülő feldolgozását. A cél egy olyan adatbázis kialakítása, amelyben személynevek, funkciók, szakirodalmi fogalmak, témakörök szerint lehet kutatni. Elegendő csupán ezek megnevezése a visszakereséshez. Eredményeként olyan forrásanyagot szolgáltat az adatbázis, amelyben a lekérdező minden igényt kielégítő információkat talál.
- *A magyar filmszakirodalom idegen nyelvű adatbázisa:* 1995. április 15. óta működik a Magyar Nemzeti Filmarchívumban. Célja a magyar filmről írott idegen nyelvű publikációk teljes körének forrásfeldolgozása és adattárolása, szolgáltatása. A feldolgozott és bevitt adatokon kívül az adatbázis a teljes (feldolgozott) idegen nyelvű publikációkat is tartalmazza. Elsősorban külföldi, tehát nem magyar nyelvű felhasználók számára készül, de igénybe vehetik a magyar film iránt érdeklődő, idegen nyelvet beszélő magyar kutatók, oktatók, egyetemi hallgatók, tömegkommunikációs szakemberek, érdeklődők is. Felépítése, adatstruktúrája, a feldolgozás elvei megfelelnek a magyar nyelvű filmszakirodalom feldolgozási szempontjainak, de tárgyszókatalógus nincs benne.

#### **Budapesti Közgazdaság-tudományi Egyetem\***

A közgazdaság-tudomány területén tölti be az országos szakkönyvtár szerepét. Online katalógusa egyben szakirodalmi adatbázisként funkcionál, amely a könyvek mellett feldolgozza a szakterület folyóiratcikkeit is. A tartalmi feltárás eléggé differenciált, kulcsszavak, szabályozott tárgyszavak, osztályozási jelzetek (kód és szöveges megnevezés) alapján lehet keresni, és a tételek némelyike referátumot (kivonat) is tartalmaz.

\*Időközben az egyetem neve Budapesti Corvinus Egyetem lett. (A szerk.)

Az országos szakkönyvtárak adatbázisai mellett vannak természetesen más, a fehér foltokat betöltő, nagyon színvonalas és érdekes adatbázisok is a piacon:

- A *MANCI (Magyar és Nemzetközi Cikkek)* adatbázis a magyar könyvtár- és információtudományi periodikumok teljes körű, kurrens és folyamatos retrospektív feldolgozása alapján épül, a külföldi folyóiratokat pedig válogatva, 1986 óta dolgozza fel. Előállítója és forgalmazója a Könyvtári Intézet Szakkönyvtára. A tételek száma jelenleg kb. 40 000, mintegy 3000 tételnyi éves gyarapodással. Az adatbázist negyedévenként frissítik. A keresés az alábbi adatelemek szerint lehetséges: szerző, kulcsszó, tárgyszó (tezasz-deszkriptor), megjelenési év, városnév, ország, régió, nyelv. A külföldi cikkek leírása referátummal is kiegészül.
- Az *ELKE* adatbázis három könyvtár számítógépes állományainak uniója, amely a kertészet, élelmiszeripar, kertépítészet, tájvédelem és tájrendezés témaköreit öleli fel. Monográfiák, szabadalmak, szabványok, újtjelentések, térképek, szakdolgozatok, folyóiratok, cikkek, előadások stb. bibliográfiai rekordjait tartalmazza. A tartalmi feltárás eszközei: ETO-jelzetek, tárgykörök és tárgyszavak.
- A Magyar Zenei Információs Központ *Musica* adatbázisa négy nyelvű adatbázis. Egy kórusmű egy rekord, s valamennyi rekord 85 mezőt tartalmaz, amely egy-egy kórusmű rendkívül kielégítő és pontos adatbevételét jelenti. A *Musica* szoftverét egy francia cég (Logi) az International Center for Choral Music megbízásából fejlesztette ki. A program különlegessége, hogy a zenei témafejt „bezongorázható”, kottaképe megjelenik, sőt meg is szólal. Így ha a karnagy csak a kezdődallamra emlékszik, a számítógép rátalál a szerzőre és a mű címére. A kórusművek között 80 szempont szerint lehet keresni, de e szempontok kombinációjával a keresés kritériumai szinte végtelenek.

A magyar tudományos szakirodalom tehát a következő tudományterületeken van számítógépes adatbázisban feldolgozva: *jogtudomány, politika, szociológia, közgazdaság-tudomány, neveléstudomány, statisztika, mezőgazdaság, orvostudomány, műszaki tudományok, nemzetiségekkel kapcsolatos irodalom, filmszakirodalom, kertészeti szakirodalom, könyvtártudományi szakcikkek*. A kórusműveket feldolgozó adatbázis különleges, főként abból a szempontból, hogy a „zenei szöveg” sajátosságaihoz igazodik.

Mind az intézményrendszer, mind a hozzá kapcsolódó szakirodalmi szolgáltatások jól tükrözik a magyarországi jellemzőket. Nincs akkora differenciáltság, mint a nemzetközi piacon, de a hazai kereslet és a szakirodalom nem is indokolja ezt. Örvendek, hogy egyre több dokumentumtípusra kiterjed a feltárás, a fontosabb tudományterületeken már hordozótól függetlenül hozzáférhető az információk. Az elektronikusan elérhető adatok többnyire a könyvtárak OPAC-jain belül kereshetők, de gyakran elkülöníthetők egymástól. A kognitívumok minden esetben a dokumentumok, szövegek fizikai határaihoz kötődnek.

A feltárás mélysége az egyes adatbázisokban nagyjából azonos. Messze nem közelíti meg a professzionális nemzetközi adatbázisok (pl. a PsycInfo) feltárási színvonalát. A szurrogátumokból sajnos elég keveset tudunk meg a szöveg tartalmáról. Maguk az adatbázisok azonban nem tartalmaznak olyan nagy rekordmennyiséget, hogy ezzel a feltárási szinttel ne lehetne kezelhető mennyiségű találathoz jutni. Az alkalmazott eszközök kicsit részletesebb vizsgálata további érdekes következtetésekre vezethet.

A magyarországi adatbázisok tartalmi feltáró rendszerét ugyanazzal a módszerrel célszerű elemezni, amellyel *1997-ben 50 nemzetközi adatbázis* feltárását vizsgáltuk. Ebben a vizsgálatban figyelembe vettünk minden olyan eszközt, amelyet az adatbázis a tartalom reprezentálására használ. A vizsgálat célja az volt, hogy minél több információt szerezzünk a tartalmi feltárás professzionális gyakorlatáról.

Számba kell venni *minden adatelemet*, amely valamilyen formában utal az eredeti szöveg tartalmára, kivéve azokat, amelyek csupán az azonosítáért felelősek (szerző, megjelenési adatok, ISBN stb.). Tehát figyelembe veendő mindenfajta osztályozási jelzet, tárgyszavazási eljárás, kiemelt tartalmi jellemző (dokumentumtípus, felhasználói célcsoport stb.), a súlyozás, a szintaxis minden formája. Ha ugyanaz az adatelem *többféle formában* is megjelenik a rekordban (pl. természetes nyelven és kóddal), akkor ez két különféle feltáró eszköznek tekintendő, amennyiben a kódolásban rejlő rendszer hozzáadott információt jelent. A számbavételen kívül vizsgálható a feltáró eszközök több tulajdonsága: mennyire *szabályozottak* (szabad szöveg vagy szabályozott jegyzék), *milyen formában* fejezik ki a tartalmat (kód vagy természetes nyelv), mennyire generikusak vagy spe-

cifikusak, milyen módon képesek tükrözni az *eredeti szövegösszefüggést*. Mindegyik jellemző valamilyen módon hozzájárul a szöveg tartalmának hívebb reprezentálásához.

A fenti szempontrendszernek megfelelően 16 magyar szakirodalmi adatbázist teszünk górcső alá.

### 1. szempont:

#### A tartalmi feltáró eszközök száma

Adatbázis	Tartalmi feltáró eszközök száma
CIKK-MMR	4
MOB	4
TEXTPAC	3
Magyar jogi szakirodalmi adatbázis	3
PRESSDOK – HUNDOK – VPI	5
MAB	5
Nemzetiségi adatbázis	5
PAD	4
Szociológiai információ	3
Filmszakirodalmi adatbázis	4
Közgazdasági adatbázis	5
ELKE	4
MANCI	3

Átlagosan négyféle eszközt használnak az adatbázisok (illetve azok szabadon hozzáférhető internetes változatai) a szövegjelentés reprezentálására. Csak az összevetés kedvéért érdemes megemlíteni, hogy a nemzetközi vizsgálatban ez az átlag 10 volt, igaz, hogy az ott vizsgált adatbázisok zöme nem szabadon hozzáférhető, csak drága előfizetési díj ellenében használható.

Nem jellemző az erősen szegmentált adatszerkezet, az is előfordul, hogy az adatszerkezetben meglévő strukturáltságot a keresésben nem lehet kihasználni. Az adatbázisok vizsgálata azt is mutatja, hogy nem megfelelőek a leírások, a kereső nagyon nehezen tudja kideríteni, hogy melyik adatelem mit takar, mi a funkciója a szövegfeltárásban, és az is ritkán derül ki, hogy melyik keresőelemet hogyan kell használni a keresésben. Ezek a hiányosságok nehezítik a hatékony keresést, így sajnos rontják a feltárás minőségét is.

Alapvető elvárás lenne, hogy a távoli felhasználó közvetlen könyvtárosi segítség nélkül is el tudjon igazodni az adatszerkezetben, és fel tudja építeni saját keresési stratégiáját. Ezen a területen a magyar adatbázisoknak mindenképpen sürgős lépéseket kell tenniük.

### 2. szempont: A tartalmi elemek jelölése

Adatbázis	Természetes nyelv	Kód
CIKK-MMR	3	1
MOB	3	1
TEXTPAC	2	1
Magyar jogi szakirodalmi adatbázis	4	0
PRESSDOK – HUNDOK – VPI	4	1
MAB	3	2
Nemzetiségi adatbázis	4	1
PAD	3	1
Szociológiai információ	3	0
Filmszakirodalmi adatbázis	4	0
Közgazdasági adatbázis	5	1
ELKE	3	1
MANCI	3	0

A feltáró eszközök között dominálnak a természetes nyelvhez kötődők. Az adatbázisok egy része alkalmaz valamilyen osztályozási rendszert, amely jelölésében (kód) is tükrözi a rendszert, a klasszikus hierarchikus rendszerek közül azonban nagyon keveset találunk. A tendencia egyrészt illeszkedik a nemzetközi trendekhez, vagyis a szakirodalmi adatbázisok mindenhol a világon – így nálunk is – saját tudományterületi osztályozásokat használnak, és nem a könyvtárakban egyébként megszokott osztályozásokat (pl. ETO). A hazai adatbázisok között is csak egy-kettő alkalmazza már az ETO-jelzeteket. Ugyanakkor azonban eltér a magyarországi trend a nemzetközitől a tekintetben, hogy a nemzetközi adatbázisok kedvelik a természetes nyelvi és a kóddal jelölt elemek együttes alkalmazását, mivel ezzel többféle megközelítést adnak a témáknak. Ennek a tendenciának még csak a csirái vannak meg a magyar adatbázisokban (pl. a közgazdaság-tudományi adatbázisnál).

### 3. szempont: A feltáró eszközök szabályozottsága

Adatbázis	Kulcs- szó	Szabad tárgy- szó	Szabá- lyozott jegyzék	Referá- tum
CIKK-MMR	1	1	2	0
MOB	1	0	2	1
TEXTPAC	1	1	1	0
Magyar jogi szakirodalmi adatbázis	1	0	2	1
PRESSDOK – HUNDOK – VPI	1	0	2	1
MAB	1	0	3	0
Nemzetiségi adatbázis	1	0	2	0
PAD	1	0	2	1
Szociológiai információ	1	0	2	1
Filmszakirodalmi adatbázis	1	3	0	0
Közgazdasági adatbázis	1	0	3	1
ELKE	1	0	3	0
MANCI	1	0	1	1

A szabályozottság a tudományos szakirodalom tartalmi feltárásában kulcsfontosságú. A keresés pontossága nagyban függ attól, mennyire biztos a feltárás egységessége. Az adatbázisok többsége alkalmaz szabályozott feltáró eszközöket, sokszor többet is, ez nagyon jó tendencia. Kevésbé jó, hogy a szabad szöveges feltárás túlnyomó részben a címre épít, nagyon kevés a szöveges tömörítés, a referátum. A szövegfeltárás szempontjából különösen nagy jelentősége lenne ennek, hiszen a mélyebb tartalom, az üzenet szinte csak itt tárható fel.

Amelyik adatbázisban referátumot is találunk, az többnyire az idegen nyelvű publikációk megértését segíti, magyar nyelvű szövegeknél még rosszabb a helyzet. Sajátos magyar nyomorúság, hogy a felhasználói igények inkább igénylik az idegen nyelvű szövegek magyar összefoglalását, mint a magyar szövegeket, mivel ez a használók nyelvtudásbeli hiányosságait hivatott kiegyenlíteni. A piac szerepe itt rendkívül erős, ez látszik például a PRESSDOK és HUNDOK adatbázisoknál, amelyek nagy hangsúlyt helyeznek a referátumokra.

A tendenciák azt mutatják, hogy a magyar adatbázisok nagyobb súlyt fektetnek a keresési igények kielégítésére, mint a szövegprezentációra. Mivel jelenleg körülbelül az utolsó 10-15 év szakirodalmi termése érhető csak el elektronikus adatbázisban, ez nem okoz problémát. Ha azonban növekedni fog az adatbázisok rekordszáma, egyre inkább gondolni kell a minőségi válogatást segítő szolgáltatásokra is.

### 4. szempont: Generikusság – specifikusság

Adatbázis	Generikus		Specifikus		For- mai tárgy- szó	Ki- emelt elem
	osztá- lyozás	tárgy- szó	osztá- lyozás	tárgy- szó		
CIKK-MMR	1			2		
MOB	1			1		
TEXTPAC				1		1
Magyar jogi szakirodalmi adatbázis				1	1	
PRESSDOK – HUNDOK- VPI	1			1		
MAB	1		1	1		
Nemzetiségi adatbázis	1			1		2
PAD	1			1	1	1
Szociológiai in- formáció	1			1	1	
Filmszakirodalmi adatbázis	1			1		2
Közgazdasági adatbázis	1			1	1	
ELKE	1		1	1		
MANCI				1		

A generikus és a specifikus feltáró eszközök egymás melletti alkalmazása szintén nemzetközi tendencia, amely megjelenik a hazai adatbázisokban is. Fontos szerepe van ennek az egészséges egyensúlynak, hiszen a generikus feltárásoknak hangsúlyos feladatuk a szövegvilág reprezentálása. Ezen a téren tehát jól állunk.

Kevésbé jó azonban a helyzet a kiemelt tartalmi elemek alkalmazásánál. Míg nemzetközileg egyre inkább uralkodó feltáró eszközzé válnak ezek a szempontok, nálunk még csak elvétve találkozunk velük. Valószínűleg a régi könyvtári gyakorlathoz való ragaszkodás az, ami nem engedi az adatbá-

zisos tervezőinek, hogy szabadabban engedjék el fantáziájukat, és külön adatmezőkben tegyenek kereshetővé tudományterületük szempontjából fontos ismérveket. Ide tartoznának a szövegek formai jellemzőire, a dokumentumtípusra vonatkozó információk is, amelyekkel szintén csak kevés helyen találkozunk.

##### 5. szempont: Relációk, kontextus kifejezése

Ez a fontos szempont annyira nincs jelen a magyar adatbázisokban, hogy már nem is lehet táblázatban ábrázolni. Arról van szó, hogy mennyiben valósul meg az eredeti szövegösszefüggés megőrzése a reprezentációban. Ahol van referátum vagy kivonat, ott van erre valami esély, egyéb esetekben azonban nem látszik, hogy a tárgyszavazást szintaktikai megfontolások befolyásolják. Az esetek többségében csak egymás után álló tárgyszavakat találunk. Az adatbázisok egy része kedveli az összetettebb szerkezeteket, amelyek rejtenek magukban valamiféle szintaxist, az azonban nem deríthető ki, hogy ez függ-e a szövegösszefüggéstől vagy sem. A tárgyszavak kombinációja, szerepoperátorok, fazetták alkalmazása egyáltalán nem jellemző.

Az Országos Pedagógiai Könyvtár és Múzeum egy évtizeden át (1989–2000) alkalmazta a PRECIS indexelési eljárást az aktuális szövegösszefüggés kifejezésére. Egy évtized tapasztalatai arra vezettek, hogy a szakirodalom jelenleg még nem feltétlenül igényli azt a mélységű feltárást, amit a PRECIS nyújt, de valamiféle szintaxisra mindenképpen szükség van az információkeresés pontossága érdekében. Ezért a PRECIS-ből megőrződött az operátorok sorrendjéből származó szintaxis, valamint az a gondolkodásmód, amelyet a PRECIS indexelést gyakorló könyvtárosok rutinszinten elsajátítottak, nagyon jól tesz az indextételeknek. Az elv tulajdonképpen annyi, hogy a tárgyszavak sorrendje közelítsen a tartalmat összefoglaló természetes nyelvű mondat sorrendjéhez.

Általánosságban megállapítható tehát, hogy a magyarországi szakirodalmi adatbázisok jelenlegi feltáró rendszere nem elegendő a részletesebb szövegrepresentáláshoz. A tudományos szakirodalom megérdemelné, hogy igényesebb szolgáltatásokat nyújtó adatbázisok álljanak a publikációk mögött, amelyek kitermelnének talán egy igényesebb információkereső közönséget is. Úgy gondolom, a színvonalas információs szolgáltatások megtermékenyítőleg hatnának a szakirodalomra is, hiszen lehetőség lenne a minőségi válogatásra.

Jelenleg ezt a válogatást az adatbázisok nem támogatják.

A minőség kulcsa véleményem szerint a befektett munkában rejlik. Az információkhoz való könnyű és gyors hozzáférés már megoldott, hihetetlenül modern technológia áll rendelkezésünkre, mégis egyre nagyobb probléma az információkeresés. A megoldást a minőségi paradigmaváltás jelentheti. A „mindenkinek mindenhol mindent” elvet hamarosan fel kell váltania a „mindenkinek mindig minőséget” elvnek. Az információs szakemberek életében ennek egyik vetülete a minőségi tartalomelemzés problémájának megoldása lesz.

A szövegfeldolgozás jelenlegi magyarországi állapotának felmérése megerősíti azt a meggyőződésemet, hogy ha a könyvtárak nem változtatnak nagyon gyorsan a tartalmi feltárás gyakorlatán, végleg lemaradnak, vagyis aláírják a halálos ítéletüket. Pillanatokon belül megszületnek azok az intézmények, amelyek hajlandók lesznek pénzt és energiát áldozni, hiszen a tudástársadalomban az értékelt és rendszerezett ismereteknek anyagi értékük van.

##### Irodalom

- CSÍK Tibor–VARGA Katalin: Knowledge and information processing. = Library automation in transitional societies. Lessons from Eastern Europe. Ed. by Andrew Lass and Richard E. Quandt. New York: Oxford Univ. Press, 2000. p. 293–312.
- MURÁNYI Péter: Az időszaki kiadványok cikkeinek feldolgozása a magyar és a lett nemzeti bibliográfiai adatbázisokban. = Könyvtári Figyelő, 3. sz. 2002. p. 511–526.
- MAGYAR Katalin: Tárgyszavazási gyakorlat és a tárgyszavak kereshetősége az Interneten elérhető könyvtári állományoknál. Budapest: Neumann Kht., 1999. <http://www.neumann-haz.hu/digital/studies/targyszol/index.html>

Beérkezett: 2005. V. 12-én.



##### Varga Katalin

az Országos Pedagógiai Könyvtár és Múzeum könyvtárának vezetője, főosztályvezető.

A Pécsi Tudományegyetem könyvtártudományi tanszékének egyetemi adjunktusa.

E-mail: [kvarga@hu.inter.net](mailto:kvarga@hu.inter.net)