

Osztályozási jelzetek és tárgyszavak konvertálhatóságának lehetőségei és korlátai

A nemzeti könyvtárban elkészült egy olyan program, amelynek segítségével a katalógus régi rekordjaiban az ETO-jelzeteket némi intellektuális segítséggel tárgyszavakra alakítják majd át. Eredményeképpen a retrospektív állományban a rekordok mind a régi jelzetek, mind pedig az OSZK-tezaurusz/Köztaurusz kifejezéseivel kereshetők lesznek. A konverzióhoz a Relex tezauruszkezelővel karbantartott, szabadon hozzáférhető OSZK-tezaurusz/Köztauruszt használják fel. A Relexben tárolt tezaurusz ugyanis jelentős részben tartalmazza már a deskriptorok és az ETO-jelzetek, továbbá az ETO-mutatószavak és az ETO-jelzetek közötti kapcsolatokat. A Java alapú program ezeket a kapcsolatokat felhasználva végzi el a konverziót. Így a retrospektív katalógusállományokban egykor ETO-val osztályozott rekordok ETO-jelzetei automatikusan, nagyobb szellemi ráfordítás nélkül konvertálhatók tárgyszavakra. Az írás Palik Máriának az e számban közölt „A tartalmi feltárás problémái online könyvtári katalógusokban” c. munkájához kapcsolódik.

Bevezető

Az Országos Széchényi Könyvtárban 2004-ben két egymást kiegészítő kísérleti projektet valósítottunk meg a tartalmi feltárás területén. A nemzeti könyvtárban korábban osztályozásra az Egyetemes Tizedes Osztályozást alkalmazták. 2004 januárjától az ETO megtartása mellett természetes nyelvű feltáró eszközként az OSZK-tezaurusz/Köztauruszt¹ vezettük be, és a Törzsgyűjteménybe kerülő dokumentumok esetén kialakítottuk a két tartalmi feltáró eszköz közötti „munkamegosztást”, összhangot.

Ez a munkafolyamat önálló tanulmányt érdemelne, mivel rendkívül hasznos, és az ETO-ra nézve pozitív adalékokkal szolgálhat az ETO kontra tárgyszavazás kérdéskörben.² Jelen írás azonban csak annak a munkának az ismertetésére vállalkozik, amelynek során megvizsgáltuk az ETO-jelzetek tárgyszavakra (tárgyi kifejezésekre) való átfordíthatóságának automatizált lehetőségét, illetve ennek a folyamatnak az inverzét, a tárgyszavakból való ETO-jelzet létrehozását automatizált eljárásban.

A kétféle tartalmi feltáró technika (osztályozás, illetve tárgyszavazás) filozófiailag eltérő szemléleten alapul, nem állítható szembe egymással, ellenkezőleg, egymást kiegészítő viszonyban van. Leegyszerűsítve a kérdést azt mondhatjuk, hogy az OSZK-tezaurusz/Köztauruszon alapuló tárgyszava-

vazás az információkereső nyelvek mellérendelő használatához tartozik, amellyel mintegy leírjuk, megfogalmazzuk a dokumentum tartalmát, és ezek a leírások – nem pedig az egyes leírásokat alkotó egyes deskriptorok külön-külön – lesznek végül is az osztályok, amelyekbe a dokumentum beletartozik.

Az ETO-t és a hozzá hasonló hierarchikus rendszereket, mivel használatuk nem nyelvyszerű, nem leíró-megfogalmazó, hanem besoroló, osztályozási rendszereknek nevezik. Maguk az egyes ETO-jelzetek jelölik az egyes osztályokat.

Az információkereső nyelvek szavai abban különböznek az osztályozási rendszerek „szavaitól”, hogy az utóbbiak eredendő rendeltetése az osztályszerű használat, az előbbieké viszont a leíró jellegű használat, amelynek csak az eredményeképpen keletkezik osztály.³ Mindebből az következik, hogy amit tárgyszórendszer vagy tezaurusz szavaival le lehet írni, az lényegében nem lesz azonos azzal, ami az ETO-jelzettel kifejezhető.

Munkánk eredményeképpen arra a meggyőződésre jutottunk, hogy a kétféle tartalmi feltáró eljárás közötti automatikus konverzió tárgyszó-ETO irányban nem lehetséges, de ETO-tárgyszó irányban körültekintő specifikálás – és némi utólagos intellektuális beavatkozás – esetén igen, vagyis az ETO-jelzet-tárgyszó irány megengedhető redundanciával megvalósítható. A fordítottja, a tárgyszava-

vak alapján releváns ETO-jelzet automatikus megalkotása a rendelkezésre álló jelenlegi eszközökkel lehetetlen.

Mindez még az utóbbi tény ellenére is nagy szó, mert éppen annak a legfontosabb mai igénynek kielégíthetőségét bizonyítja, hogy a retrospektív katalógusállományokban ETO-val osztályozott rekordok ETO-jelzetei konvertálhatók legyenek tárgyszavakra. Más szóval a retrospektív ETO-jelzet-tárgyszó konverzió eredményeként a rekordokban a meglévő ETO-jelzetek mellett nagyobb szellemi munkaráfordítás nélkül megjelenhetnek a tárgyszavak. A konverzióhoz pedig egyetemes jellege következtében éppen a szabadon hozzáférhető OSZK-tezaurusz/Köztaurusz használható fel.

Az ETO-jelzetek és tárgyszavak konverziójának projektje

Célkitűzés

A projekt célja olyan programrendszer megvalósítása volt, amely a katalógus rekordjaiban szereplő ETO-jelzetekhez automatikusan hozzárendeli az OSZK-tezaurusz/Köztaurusz megfelelő tárgyszavait. A feladatot az OSZK Informatikai Igazgatósága és a *Scriptum Rt.* közösen végezte el. A munka során az OSZK Amicus-adatbázisában a 2004. január előtt bevitt bibliográfiai rekordokhoz tárgyszavakat kellett rendelni a rekordokban rendelkezésre álló ETO-jelzetek alapján. A gyakorlatban ez azt jelentette, hogy ha a rekord 080-as HUNMARC adatmezőjében volt ETO-jelzet, akkor az ennek megfelelő szaktárgyszót vagy szaktárgyszavakat a program beírta a 650-es adatmezőjébe. Ha a jelzet dokumentumtípus-tárgyszónak felelt meg, a 655-ös (formai tárgyszó), ha földrajzi névnek, a 651-es mezőbe íródtak be a tárgyszavak.⁴

Források

Az OSZK részéről alapvető forrásnak tekintettük az Amicus-adatbázis tárgyszótáblájának (SBJCT_HDG) tartalmát. Szintén OSZK-s „alapanyag” az OSZK-tezaurusz/Köztaurusznak a Relex tezauruszkezelő és szerkesztő rendszerben [8, 9] tárolt teljes állománya, amely a relációkat és az ETO megfeleltetést illetően bővebb volt, mint az SBJCT_HDG. Harmadik forrásként a leendő új ETO-középkadás adatbázisa szolgált.

Kiindulópont

A tárgyszavak kizárólag az OSZK által karbantartott Amicus-adatbázisban szereplő tételek közül választhatók ki. A rekordokba csak olyan kifejezések kerülhetnek, amelyek az Amicus SBJCT_HDG táblájában szerepelnek.

Ismert, hogy az OSZK-tezaurusz/Köztauruszt kezelő és építő szoftver a Relex [8]. Az Amicusban található tárgyszavakat a Relexből kihozott állományból töltöttük be korábban az Amicusba. A projekt elején az Amicus SBJCT_HDG tábláját szinkronizáltuk a Relexből kinyert állománnyal. Ezt az eljárást lényegében csak a pontosítás miatt végeztük el, valamint annak érdekében, hogy a két állomány teljes összhangját biztosítsuk.

A Relexben tárolt OSZK-tezaurusz/Köztaurusz ugyan tartalmazza az ETO-mutatót a jelzetekkel és mutatószavakkal együtt, de jelenlegi állapotában még nincs feldolgozva a teljes ETO-középkadás sem a főtáblázati számokra, sem az alosztásokra vonatkozóan. Valahol félúton tart az ETO-mutatószó-deszkriptor megfeleltetés. Mindezzel együtt a munkánkhoz kiindulópontnak tekintettük a Relex állományát.⁵

A programnak a felbontott ETO-jelzetet első lépésben az ETO-mutatószóval, majd az Amicus tárgyi mutatójában (indexében) található deszkriptorral kellett megfeleltetnie. Az érintett HUNMARC mezők tartalmát a következőképpen határoztuk meg:

650-es mező: szakkifejezések helye.

651-es mező: a földrajzi tárgyszavak helye.

655-ös mező: ide kerülnek a formai tárgyszavak, amelyek a dokumentumtípust, illetve szépirodalmi művek esetén a műfajt fejezik ki.

A 650-es mező ismétlődésével a tárgyszavakat külön mezőkbe, a generikustól a specifikus fogalmak felé haladva adjuk meg. Ezek közül mindig a kronologikus (történelmi időt kifejező) tárgyszavak állnak az utolsó helyen.

A programnak gondoskodnia kellett arról, hogy ugyanahhoz a bibliográfiai tételhez egy tárgyszó csak egyszer legyen hozzárendelve. Duplumok úgy jöhetnének létre, hogy egy-egy rekordban akár öt-hat ETO-jelzet is lehetséges, és ezek tartalmazhatják ugyanazokat az alosztásokat (pl. formai: évkönyv, földrajzi: Magyarország). Nyilván nem lehet egy rekordban öt-hat ezekből.

A szoftver jellemzői

A program Java nyelven készült. Ismeri az ETO-fő-táblázatok jelzeteit és azok kapcsolódási szabályait. Felismeri az önálló és nem önálló általános közös alosztásokat. Nem önálló speciális alosztások esetén az ETO-jelzetet (ahol lehetséges) felbontja a táblázatban önállóan szereplő egységekre. Pl.:

547.29'26.07-41 [Karbonsavak észtereinek előkészítésére szolgáló berendezésekben használt reagensek]

Ebből sorrendben az alábbi részeket készíti el:

547.29 [Egyértékű telített karbonsavak. Zsírsvavak]

54.07 [Apparátus, berendezés]

547-41 [(Szerves kémia) Reagensek]

Az osztályozási jelzet és a tezaurusz-kifejezés megfeleltetésének problémái

Az osztályozási rendszerek (szintetikus megközelítés) és az információkereső nyelvek (analitikus megközelítés) elvileg, s ennek következtében formálisan nem feleltethetők meg egymásnak. Az osztályozási jelzetekhez ugyan kapcsolhatók ETO-mutatószavak, ezek azonban csak intellektuális kiegészítésre szoruló tájékozódási pontok. Az intellektuális tájékozódást támogatják a mutatószavak hátravetett értelmezői, amelyek az információkereső nyelvekben nem jelennek meg. Például:

ETO-mutatószó	ETO-jelzet	tárgyi kifejezés (deskriptor)
agykéreg (fiziológia)	612.824	agykéreg
agykéreg (patológia)	616.831.31	agykéreg
agykéreg ingerlékenysége	612.825.1	agykéreg
agykéreg ingerlékenysége	612.825.1	ingerlékenység
agykéreg mint érző és érzékelő központ	612.825.1	agykéreg
agykéreg mint érző és érzékelő központ ingerlékenység (fiziológia)	612.825.1	érzékelés
	612.816	ingerlékenység

Mindez azzal jár, hogy adott esetben az ETO-jelzetet felváltó deskriptorok elnagyoltabban reprezentálják majd a tartalmat. Példánkban a négy különféle jelzettel reprezentált tartalmat egyetlen tárgyi kifejezés, az „agykéreg” reprezentálja. Felmerült, hogy ha ugyanahhoz a deskriptorhoz több ETO-jelzet is tartozik, akkor esetleg megfontolandó, hogy egy felsőbb szintet (pl. 612.816 esetén a

612-t) is feltüntessük, mint tárgyszót. Ezt a megoldást végül elvetettük, mert nagyon sok redundáns halmazt hozott volna létre. Inkább több specifikus kifejezést használunk egy rekordban, mint kevesebb általánost.

Az ETO-jelzetekben szereplő nevek speciális problémákat vetettek fel. Szükségesnek tűnt ezeket is a bibliográfiai rekordhoz rendelni tárgyi melléktételként. Ez azonban új HUNMARC mezőt jelentett volna, más Amicus-kódértékekkel. Végül úgy döntöttünk, hogy mivel a tárgyi melléktételek készítése eddig is a formai katalogizáló gyakorlat része volt, a személy- és testületi neveket, amelyek a jelzetekben szerepelnek, nem kell a programnak kezelnie. A földrajzi neveket viszont azonosítania kell.

Problémát jelentett, hogy azokat a jelzeteket, amelyeket pontos alakban nem talált meg a program, az általánosabb (eggyel kevesebb számjegyű) jelzetnek feleltette meg. Például: mivel 07-et nem talált az elemző, a 07-es jelzeteket 0-nak értelmezte, ami elfogadhatatlan volt.

A 07-es jelzet a korábbi ETO-kiadásokban létezett (07 Hírlapok. Újságírás. Sajtóügy). A készülő új kiadásban azonban nem lesz benne, és következésképpen a programkészítés alapjául szolgáló mutató, és Relex AB állomány sem tartalmazza. Meg kellett fontolni, hogy az elemző programot ki kellene bővíteni az ilyen esetekre, hiszen ezek az esetek nem csak ennél a projektnél jelenthetnek problémát. Az elemzőt más esetekben is lehet majd használni. Az ideális persze a teljes ETO bedolgozása lenne. A korábbi középkiadások változásainak felvétele is előrelépés lett volna, különösen, ha megfontoljuk, hogy a leendő új kiadás alapján, amely az elemző program forrását jelenti, még soha senki nem osztályozott, mert még nem jelent meg. Tehát minden létező jelzet valamely korábbi kiadás alapján készült.

Ez az átfogóbb koncepció ennek a projektnek a keretei között nem valósult meg. Az viszont elfogadhatatlan volt, hogy a 07 hiányában 0 legyen a hozzárendelt jelzet, ezért azt a korlátozott megoldást választottuk, hogy a program jelzést ad, mely jelzeteket nem tudta pontosan értelmezni.

Az 1. táblázatban olyan ETO-jelzetek példái láthatók, amelyekhez a Relexben ETO-mutatószó és tárgyszó is tartozik, és a tárgyszó az Amicusban is szerepel.

1. táblázat

ETO-jelzetek példái

ETO-jelzet	ETO-magyarázat	ETO-mutatósó	Relex-tárgyszavak Amicusban: SBJCT_HDG
392.4	<i>Udvarlás, flörtölés, eljegyzés</i>	eljegyzés udvarlási szokás	udvarlási szokás eljegyzés
392.6	<i>Nemek közötti kapcsolatok. Szerelem</i>	nemek közti kapcsolatok (népr) szerelem	szerelem
654.16	<i>Vezeték nélküli táviró. Rádiótávírás</i>	rádiótávíró (hírad) vezeték nélküli távirás (hírad)	mobiltelefon
316.472.4–053.6	<i>Személyközi, interperszonális kapcsolatok. Társadalmi hálózatok. – Fiatal személyek, serdülőkorúak, tizenévesek</i>	Interperszonális kapcsolatok (szoc) Személyközi kapcsolatok (szoc)	párkapcsolat fiatalkorú

Az ETO-jelzet, az ETO-mutatósó és a tárgyszó megfeleltetése

A Relex tezauszkezelő a CF (csúcsgoalma), CA (csúcseleme) kölcsönös relációt alkalmazza az ETO-jelzet és a mutatósó viszonyának kifejezésére. A program elkészítésekor a CF/CA, az L (lásd), az L& (lásd ÉS), az LV (lásd VAGY) kapcsolatokat figyelembe kellett venni a következők szerint:

Abel-test

CF 511.236

Egyszerű helyzet: az 511.236 ETO-jelzet helyett az „Abel-test” tárgyszót kellett alkalmazni.

agyagCF 552.52
553.61
666.32
691.261

Egyetlen tárgyszóhoz több ETO-jelzet is tartozhat. Ebben az esetben – a különböző jelzetek esetén – ugyanazt a tárgyszót kellett a rekordban alkalmazni.

Abel-sokaságCF 512.742
L algebrai sokaság

Az 512.742 jelzetnek megfelelő „Abel-sokaság” kifejezés a tezauszban nemdeszkriptor, vagyis osztályozásra nem használható, ezért az L (lásd) utalásban kijelölt „algebrai sokaság” deszkriptort kellett alkalmazni.

aklimatizálódásCF 572.021.1
581.522.4
591.152
613.11
L alkalmazkodás

Ez a példa az előző kettőnek a kombinációja. Vagyis több jelzet egyetlen nemdeszkriptorhoz tartozik (aklimatizálódás), amely az „alkalmazkodás” deszkriptorra utal. Tehát az 572.021.1, az 581.522.4, 591.152, 613.11 helyett mindenhol az „alkalmazkodás” tárgyszót kellett alkalmazni.

adatgyűjtés előkészítéseCF 311.211
L& adatgyűjtés
előkészítés
statisztikai adat

A 311.211 jelzetnek megfelelő „adatgyűjtés előkészítése” kifejezés a tezauszban ún. nemdeszkriptor, vagyis osztályozásra nem használható, ezért az L& (lásd és) utalásban kijelölt összes deszkriptort kellett alkalmazni, vagyis három önálló tárgyszót, három mezőben: „adatgyűjtés”, „előkészítés”, „statisztikai adat”.

alagcsővezésCF 626.862
631.62
L& talajvíz
vízmentesítés

Két jelzet tartozik az „alagcsővezés” nemdeszkriptorhoz, amely a „talajvíz” és „vízmentesítés”

deszkriptorok együttes használatával van kifejezve. Vagyis mindkét ETO-jelzet esetén ezt a két tárgyszót kellett alkalmazni.

alak

CF	57.063.7
LV	alakok elmélete alakzat függvény

Ebben az esetben megállt a számítástechnikai tudomány; tudniillik arról van szó, hogy az 57.063.7 jelzethez tartozó „alak” nemdeszkriptor helyett az „alakok elmélete”, vagy az „alakzat”, vagy a „függvény” deszkriptorokat kell alkalmazni. Ezt pedig csak intellektuálisan lehet eldönteni. E célból ezeket külön gyűjtötte a program.

A tárgyszavak ETO-jelzetté való automatikus átfordíthatósága

Az alábbi példák azt bizonyítják, hogy egy adott dokumentumot tartalmilag feltáró, egymással szintaktikai összefüggésben nem álló tárgyszavak halmazából nem lehetséges egyértelműen azt a deszkriptorláncot (tárgyszóláncot, „mondatot”) megformálni, amely fogalmilag megfeleltethető a dokumentum tartalmának, és alapjául szolgálhatna az ETO-osztályba való besorolásnak, azaz a tárgyszó jelzetté való konvertálásának. Az egy rekordhoz kapcsolt tárgyszavak összességét nevezzük deszkriptorláncnak. A releváns ETO-jelzet létrehozásához elengedhetetlenül szükséges azon szemantikailag pontos, nyelvtanilag egyértelműen megformált deszkriptorláncnak a megléte, amely kiindulópontját képezheti az ETO-jelzetnek.

Jelenleg az Amicusban nem fejezhető ki a tárgyszavak közötti kapcsolatok a deszkriptorláncban belül (például a „Franciaország, „autó”, „export” deszkriptorlánc esetén nem tudható, hogy Franciaország által, vagy Franciaországba exportált autóról van-e szó), azaz nincs egyértelműen meghatározva a tárgyszavak közötti mondattani összefüggés. Ennek következtében több olyan önmagában helytálló ETO-jelzet, illetve jelzettelalkotható egy-egy deszkriptorláncból, amelyeknek tartalma egymástól lényeges pontokon eltér, s ezért redundáns, félrevezető, hamis jelzetalkotáshoz vezethet. A helyes mondat felismerése jelen ismereteink szerint gépi módszerrel nem lehetséges, csak intellektuális úton.

1. példa

Keresőkérdés: Magyarország ES középkor

Képes krónika. – Budapest : Nemz. Kincseinkért Egyes., 2003.

Tárgyszavak: magyar történelem
honfoglalás előtti kor
honfoglalás
kora középkor
latin irodalom
Magyarország
krónika. doktip
hasonmás kiadás. doktip
kétnyelvű dokumentum. doktip

Oszt. jelzetek: 7
943.9 „.../13” (093.3)
871-94 (439).02 =945.11
091.5 .07

Lehetséges mondatok:

1. Magyarországi latin–magyar nyelvű krónika a magyar történelem honfoglalás előtti, honfoglalás-kori és kora középkori korszakáról. Konvertálható ETO-jelzet: a rekordban szereplő három jelzet együttese.
2. Kétnyelvű Magyar történelmi krónika a magyarországi latin irodalom honfoglalás előtti, honfoglalás-kori és kora középkori korszakáról. Konvertálható ETO-jelzet: (nincs értelme, mert történelmi krónika nem szól irodalomról, de ha feltesszük, hogy mégis) 871 (439) (091) ".../13"

2. példa

Millenniumi műemlék-helyreállítási programok : a Nemzeti Kulturális Örökség Minisztériuma és az Országos Műemlékvédelmi Hivatal támogatásával helyreállított középkori templomok. – [Budapest] : NKÖM OMVH, [2003].

Tárgyszavak: műemlékvédelem
templom
középkor
Magyarország

Oszt. jelzetek: 72 .025.3/.4 (439)
726.54 (439)

Lehetséges mondatok:

1. A középkori magyarországi templomok műemlékvédelme. Konvertálható ETO-jelzet: a rekordban szereplő jelzetek együttese.
2. Templomok műemlékvédelme Magyarországon a középkorban. Konvertálható ETO-jelzetek csak időalozással bővülnek: 72.025.3/.4 (439) "10/15" 726.54 (439) "10/15"

3. példa

Keresőkérdés: művelődéstörténet ÉS reneszánsz

The panorama of the Renaissance (magyar)
Körkép a reneszánszról / szerk. Margaret Aston.;
ford. Toderó Anna. – Budapest : M. Kvk. 2003.

Tárgyszavak: reneszánsz
művelődéstörténet
művészet
15. század
16. század
Európa

Oszt. jelzetek: 7
930.85 (4) "14/15" (084.12)
7 .034 (4)

Lehetséges mondatok:

1. A 15., 16. századi európai reneszánsz művészet művelődéstörténeti tárgyalásban. Konvertálható ETO-jelzet: a rekordban szereplő jelzetek együttese.
2. A 15., 16. századi művészet reneszánsz művelődésügyi tárgyalása Európában. Konvertálható ETO-jelzet: 351.85 (4) "15/16" 7.034 (4) "15/16"

4. példa

Keresőkérdés: művelődéstörténet ÉS latin

A halhatatlan Maecenas / Cser Lászlóné. – [Pécs] :
Pro Pannonia, 2004.

(Pannónia könyvek, 0237-4277)

Tárgyszavak: Maecenas, Caius Cilnius.
művelődéstörténet
politikus
latin irodalom története
1. évezred (Kr. e.)
Római Birodalom
életrajz

Oszt. jelzetek: 7
930.85 (37) (092)
Maecenas,_C._C.
937 (092) Maecenas,_C._C.
7 .078 (37) (092) Maecenas,_C._C.
871 (091) "-00"

Lehetséges mondatok:

1. Maecenas politikus életrajza a Római Birodalom Kr. e. 1. évezredében. A latin irodalom művelődéstörténete. *(Nem sikerült egy mondatban kifejezni. Ettől kezdve pedig lehetetlen megállapítani az összetartozó fogalmakat.)* ETO-jelzetek: 937 (092) Maecenas,_C._C. 871 (091) "-00" az irodalomnak nincs művelődéstörténete, a latin irodalom története és művelődéstörténet két külön dolgot takar. 930.85 az irodalom része lehet a művelődéstörténetnek.

Példák automatikusan eldönthetetlen többértelműségekre

1. tetőtér-beépítés

ETO-jelzete: 69.059.38

OSZK-tezauruszban: **tetőtér-beépítés**L& épületbővítés
padlás

Vissza az ETO-ba: 69.059.38

643.9 [padlás: irreleváns]

2. társasági törvény

ETO-jelzete: 347.72

OSZK-tezauruszban: **társasági törvény**L& társas vállalkozás
törvény

Vissza az ETO-ba: 658.114.2 [társas vállalkozás]

094.5 [egyres törvények]

Az OSZK tárgyszavazási gyakorlatának a gazdasági társaság + Magyarország + jogszabálygyűjtemény felel meg.

3. Jézus élete

ETO-jelzete: 232.9

OSZK-tezauruszban: Jézus

életrajz

Vissza az ETO-ba: 232 [Jézus]

(092) [életrajz]

Mivel Jézus életéről van önálló jelzet, értelmetlen lenne a visszafordított jelzetekkel osztályozni.

4. joga

ETO-jelzete: 159.962.7 [joga, pszichológia]

294.527 [joga, vallás]

613.71 [joga, egészségügy]

615.851.86 [joga, terápia]

A „joga” tárgyszó konvertálása négyféle eredménnyel járna.

5. család

ETO-jelzete: 314.6 [család, háztartás,
demográfia]

316.356.2 [család, szociológia]

A „család” tárgyszó konvertálása kétféle eredménnyel járna.

6. szeretet

ETO-jelzete: 172.15 [hazafiság hazaszeretet]

177.61 [szeretet,_etika]

177.7 [emberszeretet]

241.513 [szeretet,_morális_
teológia]

241.513.2 [felebaráti szeretet]

316.647.2 [empátia szeretet
szerelem rokonszenv barátság]

A „szeretet” tárgyszó konvertálása hatféle eredménnyel járna.

Összegezés

A két projekt tapasztalatai alapján az alábbi következtetéseket szűrtük le:

- A két osztályozási rendszer filozófiájának különbözősége és a fent vázolt praktikus nehézségek miatt információvesztés, zaj nélküli pontos megfeleltetés automatizált úton lehetetlen. Mérlegelni kellett, hogy a konverziós folyamat milyen eredménnyel jár.
- Megítélésünk szerint az ETO–tárgyszó irányban végzett konverzió eredményeképpen az információ torzulása nem olyan mértékű, hogy az eredeti tartalmat jelentősen meghamisítaná.
- A kereshetőség hatványozottan javul, ha egy dokumentumot ETO-val is és tárgyszóval is osztályoznak.
- Ezért az ETO–tárgyszó irányú konverzió megfelelő körütekintés és előkészítés utáni elvégzése nyereség.
- A tárgyszó–ETO konverzió automatizáltan még jelentős intellektuális ráfordítás (javítás) esetében sem hatékony.

Mindkét projektben és ezen írás létrejöttében is nagy segítségemre volt Szalóki Gabriella és Ungváry Rudolf. Munkájukat ezúton is köszönöm.

Jegyzetek

- ¹ Az átfogó magyar teaurusz kettős neve abból ered, hogy egyrészt a nemzeti könyvtár, másrészt pedig a közművelődési könyvtárak teaurusza, de számos más könyvtárban is használják [5, 7].
- ² A kérdéstről a Katalist könyvtári levelezőn folytatott vitát összefoglalja Fejős [2], az ETO és a deskriptoros nyelvek filozófiájának különbségéről és az ETO helyzetéről lásd még [6].
- ³ Az információkereső nyelvek szavairól (tárgyszó, tárgyi kifejezés, deskriptor, jelzet stb.) és az információkereső nyelvek és osztályozási rendszerek filozófiájáról lásd [10].
- ⁴ A HUNMARC-ról és használatáról lásd Barki és Prejczer színvonalas munkáját [1].
- ⁵ A Relex és az Amicus tárgyszóállomány összhangjának megőrzése jelenleg azért fontos, mert az Amicus még csak a nemdeskriptorok és deskriptor közötti kapcsolatokat kezeli, de nem kezeli a deskriptorok közötti kapcsolatokat. Idővel az Amicus fejlesztésében is el lehet jutni majd oda, hogy az utóbbi kapcsolat kezelése lehetővé válik, és akkor a kapcsolatok áttöltése a Relexből fog megtörténni.

Irodalom

- [1] BARKI Katalin–PREJCZER Paula; az elektronikus változatot szerkesztette Barki Katalin: Számítógépes katalogizálás Hunmarc formátumban. – Berzsenyi Dániel Főiskola Központi Könyvtár; Berzsenyi Dániel Főiskola könyvtár- és információtudományi tanszék. – Szombathely: BDFKK, 2003. <http://www.bdff.hu/konyvtar/oktatas/marc/index.htm>
- [2] FEJŐS LÁSZLÓ: Kell-e még az ETO? Vita a könyvtárosok Katalist levelező fórumán (Összegzés). Budapest: 2000. <http://www.ki.oszk.hu/3k/valcikkek/valcikkek0003/fejos.html>
- [3] HORVÁTH ZOLTÁNNÉ: Az ETO védelmében. Előterjesztés a fejlesztés és az együttműködés lehetőségeiről. = Könyvtári Figyelő, 3. (39.) köt. 3. sz. 1993. p. 373–383. http://www.ki.oszk.hu/kf/1993/3/horvathne_h.html
- [4] HORVÁTH ZOLTÁNNÉ: ETO és tárgyszó: ellentét helyett együttműködést! Haladás „megtagadás” nélkül. = Könyv, Könyvtár, Könyvtáros. 1993. május, p. 22–28.
- [5] OSZK-teaurusz/Köztaurusz. http://www.oszk.hu/frame_hu.htm?hun/szakmai/teaurusz/teaurusz_index_hu.htm
- [6] UNGVÁRY Rudolf: Az ETO szükségessége. = Iskolakultúra, 10. köt. 4. sz. 2000. p. 27–31. <http://www.iskolakultura.hu/documents/2000/4/tanulm.pdf>
- [7] UNGVÁRY Rudolf: Az OSZK teaurusz és a KÖZTAURUSZ. = Könyvtári Figyelő, 11. (47.) köt. 1. sz. 2001. p. 11–40. <http://primula.oszk.hu/kiadvany/kf/2001/1/ungvary.html>
- [8] UNGVÁRY Rudolf: Teauruszkezelő programok és a RELEX. = Tudományos és Műszaki Tájékoztatás, 48. köt. 1. sz. 2001. p. 3–16. <http://ontologia.hu/Members/ungvary/teauruszbibliografia/articlereference.2004-12-25.1608756848>
- [9] UNGVÁRY Rudolf: A teaurusz a felhasználói felületen. Az optimális teaurusz megjelenítés problémái. = Tudományos és Műszaki Tájékoztatás, 48. köt. 3. sz. 2001. p. 99–108. http://tmt.omikk.bme.hu/show_news.html?id=1620&issue_id=26
- [10] UNGVÁRY Rudolf: Az információkeresés szavai. = Tudományos és Műszaki Tájékoztatás, 50. köt. 12. sz. 2003. p. 1–27. <http://www.hik.hu/tankonyvtar/site/books/b153/ch-8.html#ftn.d4e3240>

Beérkezett: 2005. IX. 26-án.



Bánki Zsolt István

az Országos Széchényi Könyvtár Informatikai Igazgatóságának munkatársa.

Kutatási területe a természetes nyelvű információkereső nyelvek.
E-mail: bazso@oszk.hu