

Diplomatezauruszok hibái és tanulságai

Néhány diplomamunkaként készült tezauruszt elemeztünk. Rávilágítunk azokra a jellegzetes kezdeti hibákra, melyek az egyébként átlagon felül motivált szerzők műveit jellemzik. Rájtuk keresztül némi képet kaphatunk a könyvtári szakma tezauruszokkal kapcsolatos tájékozottságáról is. A diplomamunkákban olykor mellőzik a magyar szabványokat, problémák tapasztalhatók a fogalmi és a nyelvi szint megkülönböztetésében és a generikus reláció kezelésében. Mindez nem független a katalogizálás- és osztályozáselmélet hazai színvonalától. A tezauruszok oktatása elválaszthatatlan a MARC21 és a vele összhangban levő HUNMARC besorolási adatcsere-formátum oktatásától.

Bevezető

Előfordul, hogy diplomamunkaként tezauruszt készítenek. Üdvözlendő és értékes törekvésről van szó, ráadásul olyan hallgatókról, akikben az átlagosnál erősebb a szakmai motiváltság. A tanulságok miatt érdemes ezekkel a munkákkal a nyilvánosság előtt is foglalkozni, hogy szerzőik és mindazok, akik körülöttük bábáskodtak, lássák, komolyan vesszük őket, számítunk rájuk a nemzet csinosításában. Olyan gyerekbetegségekre derülhetne fény, amelyek elkerülhetők.¹

A tárgyalt tezauruszok megérdemlik a figyelmet, látszik, hogy szerzőik nemcsak jól ismerik azt a szakterületet, amelynek tezauruszát elkészítették, hanem lelkiismeretesen követték mindazt, amit a tárgykörben tanítottak nekik. Hibáik nem vonnak le semmit törekvéseik jelentőségéből.

A magyar tezauruszszabvány alkalmazása

Több diplomatezauruszban a relációk nevét az angol szabvány [4] szerint adták meg, holott létezik érvényes magyar szabvány [6]². Ennek következtében a magyar nyelvű tezauruszcikkek belül idegen nyelvű jelölések jelennek meg (1. ábra).

Nem állja meg a helyét az az érv, hogy azért célszerű angol szabványos relációjeleket [3] [4] használni, mert ezeket a magyar nyelvterületen kívül is értik. Egy magyar tezauruszban az összes lexikai egység magyarul szerepel, a magyarul nem tudó használó akkor sem tud semmit kezdeni velük, ha érti a relációt, de nem érti az utána szereplő kifejezést.

ruházat	
SN	A testet takaró, rendszerint textil anyagból való egy vagy több részes készítményt öltözetnek vagy ruhának, ruházatnak hívjuk
UF	öltözet
BT	textiláru
NT	<ul style="list-style-type: none">• alkalmi ruházat<ul style="list-style-type: none">• • gyászruha• egyenruha• népviselet
RT	öltözködés ruhakiegészítő szabóság

1. ábra Deszkriptorcikk az angol szabvány szerinti relációjelekkel

Magyar tezauruszban nem az angol, hanem az MSZ 3418 szabvány magyar relációjeleit kell alapértelmezésben használni. Ez egyébként nem zárja ki egyrészt azt, hogy a tezaurusz esetleg elkészüljön angol változatában az angol relációjeleket alkalmazzuk. Sőt, a magyar változatban alkalmazhatunk felhasználóbarát relációneveket is, de magyarul. Példa rá a *Magyar Nemzeti Filmarchívum* tezaurusza³ (2. ábra).

Ezen túlmenően a tezauruszokra is vonatkozó angol MARC21, és a vele kompatibilis magyar HUNMARC [2] besorolási adatcsere-formátum kifejezetten lehetővé teszi, hogy egyazon kezelőrendszeren belül a tezaurusz ne csak a magyar, hanem az angol, vagy akár a fenti magyar relációjelölésekkel, illetve felhasználóbarát leíró megnevezésekkel (2. ábra) jelenjék meg a nyilvános hozzáférésű online katalógusban (OPAC). A mai

könyvtári rendszerekben elvileg magától értetődő a fenti adatsere-formátumok szerinti kezelés. Más lapra tartozik, hogy az adatsere-formátumok lehetőségeire felkészített rendszerek magyarországi alkalmazása során nem használják ki a lehetőségek jelentős részét (a kérdésre még visszatérünk).

kisebbségi kérdés		a)
<i>helyettesíti vagylagosan</i>	politikai kérdés	
<i>átfogóbb</i>	nemzetiségi kérdés	
<i>speciálisabb</i>	politika	
<i>teljesebb</i>	kisebbség	
<i>része</i>	nemzetiségi szervezet	
	konfliktus	
<i>irányul</i>	autonómia	
<i>lásd még</i>	függetlenségi mozgalom	
	kitelepítés	
	kollektív jogok	
	nacionalizmus	
	politikai mozgalom	
	szórványmagyarság	

kisebbségi kérdés		b)
nyelvhasználat HV	nyelvhasználat	
F	politikai kérdés	
A	kisebbségi kérdés	
T	politika	
P	kisebbség	
	nemzetiségi szervezet	
R	konfliktus	
X	autonómia	
	függetlenségi mozgalom	
	kitelepítés	
	kollektív jogok	
	nacionalizmus	
	politikai mozgalom	
	szórványmagyarság	

2. ábra **A Magyar Nemzeti Filmarchívum deszkriptorcikke [1] felhasználóbarát, magyarázatos relációnevekkel (a), és szabványos relációjelekkel (b)**

A diplomatezaurusz készítője nem abból tanulja a legtöbbet, ha angol relációjeleket használva készítetik el vele a tezauruszt, hanem abból, hogy diplomatezauruszában, annak előszavában saját tezauruszában lexikai egységei alapján bemutatja, hogyan kell az ún. hivatkozásszöveg-vezérlő almezővel a tételfejek szerepét játszó deszkriptorok és nemdeszkriptorok esetében az angol vagy

a felhasználóbarát online megjelenítéseket kezelni. A 3. és a 4. ábrán az 1. ábra deszkriptorcikkét alapul véve ezt az utóbbi esetet példázzuk. A besorolási formátumokban szokásos módon, a \$ jellel jelöljük, hogy az utána következő betű almezőhívójel.

A 150-es mezőben a tételfej (esetünkben egy deszkriptor), a 450-es és az 550-es mezőkben a tételfejhez kapcsolódó kifejezések szerepelnek. A \$w utalásvezérlő almezőben megadott *y*, *g*, *h* és *m* értékek jelzik a reláció típusát (helyette, fölrendelt, egyéb rokonság), melynek alapján a kezelőrendszer az alapértelmezett magyar relációjeleket jeleníti meg. A 680-as mező a nyilvános általános megjegyzés helye. A # a mezőhívójel után az indikátorpozíciókat jelzi (esetünkben azt, hogy nincs értékük), a 040-es mező a rekord forrása, a 005-ös a rekorddal végzett utolsó művelet időpontja, a 008-as mező pedig a rekord különféle tulajdonságainak minősítését tartalmazza (a rekord létrehozásának dátuma, kitüntetett tételfej [azaz deszkriptor], tárgyi melléktételben felhasználható stb., lásd részletesen a MARC21, illetve a HUNMARC besorolási adatsere-formátumban [2]). A pozíciók könnyebb azonosíthatósága érdekében a 008-as mező ötös csoportokra való tagolását is a MARC21 szerinti | jel biztosítja. A 008-as mezőben a félkövér # jelek a HUNMARC szerinti betöltetlen (felhasználatlan) pozíciókat jelölik, a dőlt # jelek pedig minősítés-képviselő üres helyek. A pozíciók részletes magyarázatát a 10. ábrához kapcsolódva ismertetjük.

Mint arra utaltunk, angol relációjelekkel magyar tezauruszcikket megjeleníteni haszontalan dolog. Magyar tezauruszt általában nem angol anyanyelvűek használnak, hanem magyarok; angol relációjeleket angol nyelvű tezauruszban, adott esetben a kétnyelvű (magyar–angol) tezaurusz angol változatában kell használni. Ezért a 4. ábrán nem az angol relációjelekre váltást, hanem a felhasználóbarát relációmegnevezésre való átváltást mutatjuk be.

A \$w utalásvezérlő almezőben megadott \$i utalásszöveg-vezérlő almező értéke elfojtja a rendszerben alapértelmezett magyar relációjel-megjelenítést, és az \$i utalásvezérlő almezőben megadott, az ábrán a megjelenítésben dőltén írt relációjelet jeleníti meg.

Tag ii \$ nz n # 5.
 000 #####nz 22#####n 4500
 001 <rekordazonosító>
 005 20080713233723.0
 008 080713 | #n#an | z##ba | #n### | ##### | ###a# | ana## | ###d
 040 ## \$a<egyetem kódja, ahol a diplomatezaurusz készült> \$bhu \$f<a diplomatezaurusz neve>
 150 ## \$aruházat
 450 ## \$wy \$aöltözet
 550 ## \$wg \$atextiláru
 550 ## \$wh \$aalkalmi ruházat
 550 ## \$wh \$aegyenruha
 550 ## \$wh \$anépviselet
 550 ## \$wm \$aöltözködés
 550 ## \$wm \$aruházatkiegészítő
 550 ## \$wm \$aszabóság
 680 ## \$aA testet takaró, rendszerint textilanyagból készült, egy vagy több részes mesterséges burkolat

Megjelenítés magyar szabvány szerint:**ruházat**

M: A testet takaró, rendszerint textilanyagból készült, egy vagy több részes mesterséges burkolat
 H öltözet
 F textiláru
 A harántszalag
 egyenruha
 népviselet
 X öltözködés
 ruházatkiegészítő
 szabóság

3. ábra Tezauruszcikk rekordja a besorolási adatsere-formátum szerint, és megjelenítése szabványos magyar relációjelekkel

Tag ii \$ nz n # 5.
 000 #####nz 22#####n 4500
 001 <rekordazonosító>
 005 20080713233723.0
 008 080713 | #n#an | z##ba | #n### | ##### | ###a# | ana## | ###d
 040 ## \$a<egyetem kódja, ahol a diplomatezaurusz készült> \$bhu
 f <a diplomatezaurusz neve>
 150 ## \$aruházat
 450 ## \$wiy \$aöltözet \$inem használt/szinonimája
 550 ## \$wig \$atextiláru \$iátfogóbban
 550 ## \$wih \$aalkalmi ruházat \$ifajtája
 550 ## \$wih \$aegyenruha \$ifajtája
 550 ## \$wik \$anépviselet \$ifajtája
 550 ## \$wim \$aöltözködés \$ilásd még
 550 ## \$wim \$aruházatkiegészítő \$ilásd még
 550 ## \$wim \$aszabóság \$ilásd még
 680 ## \$iA testet takaró, rendszerint textilanyagból készült, egy vagy több részes mesterséges burkolat

Megjelenítés a formátumban megadott relációnevekkel:**ruházat**

Magyarázat: A testet takaró, rendszerint textilanyagból készült, egy vagy több részes mesterséges burkolat
 nemhasznált/szinonimája öltözet
 átfogóbban textiláru
 fajtája harántszalag
 egyenruha
 népviselet
 lásd még öltözködés
 ruházatkiegészítő
 szabóság

4. ábra Tezauruszcikk rekordja a besorolási adatsere-formátum szerint, és megjelenítése magyar kezelőrendszerben felhasználóbarát relációnevekkel

Mindehhez természetesen a hallgatóknak az oktatás során meg kell ismerkedniük a besorolási adatok adatsere-formátumával, amire még kitérünk.

A teaurusz főrészeinek szerkezete

A teauruszok főrészeinek minden nemzeti és nemzetközi szabványban egyforma a szerkezete, és a szabványok alapvető előírása a főrész megléte. Nem szerencsés csak olyan főrészt megjeleníteni, amelyben a teauruszcikket a szabványoktól eltérő formában jelenítik meg. Egyes teauruszokban például hierarchikusan több szinten strukturálták a főrész teauruszcikkeit (1., 5. ábra).

<p>tárgyábrázolás</p> <p>F képek</p> <p>A • állatokhoz kapcsolódó tárgyak</p> <ul style="list-style-type: none"> •• kantár •• kengyel •• méhkas •• nyereg •• ostor •• patkó •• sarkantyú •• zabla • ékszerek <ul style="list-style-type: none"> •• gyűrű • épületek <ul style="list-style-type: none"> •• bástya •• vár • épületrészek <ul style="list-style-type: none"> •• kapu <p>stb.</p>
--

5. ábra Részlet egy diplomateaurusz többszintes hierarchikus deszkriptorcikkből

Hogy a hierarchiaszinteket jelölő jobbra irányuló eltolásokat jobban kiemeljék, az alárendelt deszkriptorok előtt félkövér pontokkal jelezték a szint mélységét (• = 1. szint, •• = 2. szint stb.). Ez a jelölés a szabványok szerint nem használatos, és nem ok nélkül.

Az ilyen szerkezetű főrészek ugyanis alkalmatlanok arra, hogy a könyvtári rendszerek manuális átszerkesztés nélkül átvegyék őket, és nem konvertálhatók a teauruszokra vonatkozó nemzetközi és magyar adatsere-formátumra⁴ sem (ami egyébként a feltétele annak, hogy bármelyik mai könyvtári rendszerbe át lehessen venni). A könyvtári rendszerek eleve nincsenek felkészítve ilyesfajta kifinomult teauruszcikk-formára, csak a

Ungváry R.: Diplomateauruszok hibái és tanulságai

szabványos szerkezetű teauruszcikket képesek kezelni és megjeleníteni. Magának a pusztá generikus hierarchiának a megjelenítésére a külön hierarchikus rész való. Ez utóbbi megjelenítés használható majd fel távlatilag ahhoz is, hogy ontológiákban hasznosítsák őket. De ezt a hierarchiát, mivel a gyakorlatban a szabványos teauruszcikk alapján a mindenkorai kezelőrendszer jeleníti meg (ha egyáltalán fel van készítve erre, ami jelenleg nem jellemző), legfeljebb gyakorlás céljából célszerű a hallgatókkal elkészíttetni. A teaurusz alapja tehát mind a végfelhasználó számára történő megjelenítést, mind az informatikai kezelést illetően kizárólag a szabványos teauruszcikk (és annak besorolási adatsere-formátuma).

A mondottak igazságán nem változtat, hogy számos nemzetközi teaurusz esetében is megfigyelhető ilyen szofisztikáltabb teauruszcikk-szerkezet. Általában azért, mert olyanok készítették, akik akkor, és csak ezt az egyetlen teauruszt készítették el; egy kicsit elszállt velük a képzeletük.

Felesleges teauruszrészecskék

Előfordult, hogy egy teauruszban ugyan a magyar szabvány szerinti relációjeleket alkalmazták, de hozzá fáradságos manuális munkával kidolgoztak olyan részt, amelyre semmiféle szükség nincs. Minden lexikai egységhez azonosítót rendeltek hozzá, és erről betűrendes mutatót készítettek (6. ábra).

<p>Zabumba 470, 4190</p> <p>Zambumbia 4200</p> <p>Zarb 110, 1750, 3800, 4210</p> <p>Zeneszerzők 4220</p> <p>Zongora 1280, 2860, 2880, 4230</p>

6. ábra Felesleges azonosítók (és felesleges mutatójuk) egy diplomateauruszban

A teauruszokban egyáltalán nincs szükség mesterségesen létrehozott azonosítókra. A lexikai egységek azonosítóival a felhasználó semmi sem tud kezdeni. Számára egyértelműen azonosítva vannak a fogalmak a megnevezéseikkel, hiszen a teauruszok lényege, hogy szabványosított nyelv szótárai. Más szóval egy teauruszban nincs, nem lehetséges két egyforma megnevezésű lexikai egység. Maguk a megnevezések az elsődleges azonosítók, mivel egyértelműek és nem ismétlődnek. A nemdeszkriptor–deszkriptor kapcsolatokkal

és a hátra vetett értelmezőkkel (a szabvány szerint a kerek zárójelek közé tett kiegészítő adattal) kell megoldani a szinonimák és a homonimák szabályozását.

Például a „kutya”–„eb” szinonimapár esetében, ha deskriptorként (kitüntetett egységesített besorolási adatként) a „kutya” kifejezést választjuk, akkor a szabályozás formája (ahol a félkövéren megjelenített „kutya” a deskriptor, és az „eb” a nemdeskriptor, melyet helyette használni kell, vagy melyet a rendszer helyette automatikusan használ):

kutya

L eb

Például a „sorozat”, a „rév”, a „Füzesd” és a „forrás” eseteiben (a deskriptorokat félkövéren megjelenítve):

sorozat (matematika)
sorozat (dokumentum)

rév (hajózás)
Rév (Bihar vármegye)

Füzesd (Hunyad vármegye, Dévai járás)
Füzesd (Hunyad vármegye, Puji járás)

forrás

LV **erőforrás**
fényforrás
forrás (vízforrás)
jelforrás
történelmi forrás

A lexikai egységeknek természetesen van azonosítója, de ezzel se a felhasználónak, se a tezauszszerkesztőnek semmi dolga. Minden kezelőrendszer, amely a tezausz (az adatcsereformátum szerkezetében [2]) átveszi, a rendszer szabályai szerint automatikusan a saját folyamatos számozású azonosítóját generálja a lexikai egységekhez, és egyáltalán nem veheti figyelembe az intellektuálisan kialakított azonosítókat. Ez az azonosító azonban kizárólag az adatcsere-formátum rekordjának kezeléséhez szükséges⁵, és az se biztos, hogy a mindenkori adatbázis átrendezése esetén változatlan marad.

Kézi vagy szerkesztőprogrammal támogatott diplomatezausz-készítés?

A szakirodalomból ismert, hogy létezik Magyarországon tezauszkezelő és szerkesztő program (a Relex)⁶. Sokan mégis manuálisan készítik el dip-

lomamunkaként a tezauszt. Ez tőről metszett kínlódás, hiszen még néhány száz lexikai egységből álló tezausz esetében is nagyon gazdag relációszerkezetet kell intellektuálisan kezelni, ami hibátlanul jóformán lehetetlen.

Ennek a következménye, hogy nem küszöbölhető ki megbízhatóan az ellentmondó és hiányzó kapcsolatok. Például az egyik tezausz hierarchikus tezauszcikkében a *szarufa* a harántszalagnak van alárendelve, a *szarufa* deskriptorcikkében viszont a *szarufa* az egy szinttel magasabban levő harántosztásnak az alárendeltje.

Előfordul, hogy a tezausz végső változatát PDF-ben rögzítik. Ebben a formában azonban semmiféle szerkesztőrendszerbe (így a Relexbe) sem lehet átvinni. Ha viszont Word állományban kerül sor a megjelenítésre, és ez a megjelenítés az MSZ 3418 magyar szabványnak [6] megfelelő szerkezetű, akkor némi programozói munkával átvehető adott könyvtári rendszerbe, és minden kézi közbeavatkozás nélkül átvehető például a Relex tezausz-szerkesztőbe [14], melyben egyetlen billentyűparancsra megjelenik minden formálisan kiszűrhető hiba (mint pl. az inverz kapcsolat hiánya, a többszörös ellentmondó kapcsolat, a reflexívkapcsolat, a kör és a hurok). Amelyik diplomatezauszt az elmúlt években tezauszszabvány szerinti tezauszcikkek Word szövegállományának formájában alakították ki (nem kis fáradságos munkával), és eljuttatták e cikk szerzőjéhez, annak a tezausz-nak a formális hibáit a Relexbe átvéve gyorsan ki lehetett küszöbölni.

A fogalom fogalmának ismerete

Gyakran tapasztalható bizonytalanság a fogalmi és a nyelvi szint szétválasztásában, annak módjával. A hallgatók dolgozataikban, illetve diplomatezauszaiak előszavaiban például a fogalmak „tárolásáról” írnak.

A fogalom nem tárolható, nincs helyesírása, és se manuálisan, se automatizáltan nem kezelhető. A fogalom ugyanis csak a tudatban létezik. Tárolni, helyesen írni a fogalmak megnevezéseit, a szavakat (pontosabban a közneveket) lehet. Ha a szavak szótárban (így a strukturált szótárakban is, mint amilyen a tezausz) szerepelnek, akkor lexikai egység a nevük.⁷

Néhány példa az egyik diplomatezausz előszavából:

„A fogalom magyar és idegen nyelvű változata...”⁸

Helyesen: A lexikai egység magyar és idegen nyelvű névalakja...

„A fogalmak összegyűjtését követően formai és tartalmi egységesítést volt szükséges elvégezni, mivel a szóalakok forrásmunkában szereplő változata nem minden esetben volt alkalmas deskriptornak. Így a fogalmak alakját helyesírási, nyelvtani, lexikográfiai és szakmai szempontok szerint is rendezni kellett.”

Helyesen: A szógyűjtés után formai és tartalmi egységesítés következett, mivel a névalakok forrásmunkában szereplő változata nem minden esetben volt alkalmas lexikai egységnek. Így a kifejezések alakját helyesírási, nyelvtani, lexikográfiai és szakmai szempontok szerint is egységesíteni kellett.

Ugyancsak nem tűnik világosnak, mi a szerepe a fogalomnak a nyelvi megnevezéshez képest, és mi e kettő viszonya a dolgokhoz, előfordulásokhoz, egyedekhez (összefoglalóan a referenciákhoz), melyekre vonatkoznak. Az egyik tezaurusz előszavában szerepelt például: „...a deskriptorok... fogalmat reprezentálnak.” A deskriptorok azonban szabványosított megnevezések, melyek a fogalmakat nem reprezentálják, hanem azonosítják. Reprezentálni a fogalmak reprezentálják a referenciákat (a dolgokat). A helyes szóhasználatot az 1. táblázat tartalmazza.

A lexikai egységek egyes és a többes számú alakja, avagy a fogalmi és a terjedelmi szint

Az egyik tezaurusz előszavában olvasható:

„... külön kellett választani:

- a főnévként megadott fogalmakat, pl.: zene-mű, Nap, szarvas;
- a jelzős szerkezeteket, pl.: kerti lilium, tört arany, rangjelző korona.”

1. táblázat

A szintek viszonyai

szint	egység	szerep
fogalmi szint	fogalom/osztály	reprezentálja a referenciákat (dolgokat)
nyelvi szint	deskriptor	azonosítja a fogalmat
	nemdeskriptor	a deskriptor által helyettesített megnevezés ⁹
	jelzet	azonosítja a fogalmat/osztályt
	köznév (köznyelvi megnevezés)	megnevezi a fogalmat (pl. hegy)
	tulajdonnév (egyed neve)	megnevezi az adott dolgot, individuumot (pl. Hargita)
tárgyi szint	referencia	a megnevezés jelölete, a fogalom terjedelmének eleme (az, amit a fogalom reprezentál)

A kétféle névtípus különválasztását a tezaurusz-szerkesztő feltehetően azért tartja szükségesnek, mert az uniterm (egyetlen szóból vagy összeforrott szókapcsolatból álló) kifejezések egy részét többes számban adja meg. Valójában azonban információkereső-nyelvi szóként mindegy, hogy valami uniterm vagy jelzős szerkezetű összetett szó. Ráadásul a felsorolt jelzős szerkezetek ugyancsak főnevek, alaptalan a „főnév” versus „jelzős szerkezet” alternatíva.

Semmiféle információkereső vagy számítástechnikai érv nem szól az ilyen szétválasztás mellett (amiből nem következik, hogy ne volna szükség szétválasztásra, ezzel a következő fejezetben foglalkozunk).

A fenti példában megadott jelenség ugyanannak a szemléletnek a következménye, mint amelynek alapján a diplomatezauruszokban (sokszor a gyakorlatban használt tezauruszokban is) olykor többes számban adják meg a lexikai egységeket. Az egyik előszó szerint „többes számot alkalmaztam azoknál a főneveknél, ahol egyértelmű volt, hogy többféle alakban fordul elő. Pl.: madarak, halak, fák, égitestek”.

Mi nem fordul elő többféle alakban? A tégláknak, kavicsoknak, emberi füleknek is számtalan alakja, mérete van... Ráadásul az előszó fogalmazója valójában nem „alakra” gondolt, hanem „fajtára”. De nincs olyan fogalom, amelynek ne lennének fajtái (ez még a legspeciálisabb, legalsóbb szinten is csak egyre összetettebb szavak alkalmazásának kérdése). A többes szám szintaktikai kategória, más szóval közvetlenül állítás kifejezésére való: arra, hogy kifejezzük, egy vagy több dolog alkotja a kijelentés tárgyát. Szótárak lexikai egységei nem kijelentések, ha a szótárban az szerepel, hogy „tégla”, akkor nem a tégláról, hanem a téglát reprezentáló egyetlen fogalom nevééről van szó.

A tezaszuszuszkészítési gyakorlatban – melyet a szabványok is támogatnak – csak azokat a lexikai egységeket célszerű többes számban megjeleníteni, amelyeknek nincs egyes számú alakja, vagy megtévesztő volna az egyes számú alak (például „javak”, „közjavak”, „tőkejavak”; harkályalakúak). Kivételek ugyan lehetnek, de ritkák (pl. „évszakok szerinti változás”).

Másrészt állandó bizonytalanságot okoz (és nincs rosszabb, mint a bizonytalanság; ezt felhasználó, feldolgozó [és kisgyermekes szülő] egyaránt tudja), ha a lexikai egységek egy részét egyes számban, más részét többes számban szerepeltetik. Nem adható pontos feltétel, amely meghatározná, hogy milyen szavak esetén kellene használni. Ha minden anyag és tárgy név többes számú ('műanyagok', 'zongorák'), akkor ennek mi az előnye? Ha meg csak az általánosabb anyagok és tárgyak, akkor a specifikusságnak melyik szintjétől legyen egyes számú? A 'műanyagok' többes szám, a 'réz' pedig nem?

Szögezzük le, hogy teljesen felesleges, ha az átfogóbb fölérendeltek többes számú alakban, a speciálisabbakat pedig egyes számban adják meg. Ennek a gyakorlatnak az a vége, hogy a több ezer kifejezésen belül végül már csak teljesen önkényesen lehet eldönteni, mi álljon egyes, és mi többes számban? Például:

műanyag(ok)
 szintetikus műanyag(ok)
 polimerizátum(ok)
 polietilén(ek)
 poliolefin(ek)
 hőre lágyuló poliolefin(ek)
 termoreduktív poliolefin(ek)
 Ciba-Gegy MZ25

Mi dönti el, hogy a fenti hierarchialáncban melyik szinttől lefelé nem kell a többes szám? Minden fogalomra jellemző, hogy terjedelmébe végtelen számú egyed tartozhat. Igaz ez mind a hierarchia csúcsán szereplő műanyag fogalmára, mind pedig a hierarchia legalján szereplő Ciba-Gegy MZ25 márkanévű termoreduktív poliolefinre.

Az egyik tezaszuszuszban például a madarak fordulnak elő többféle – a szerző szerint – „alakban”. Ha a madarak többes szám, akkor a varjú miért nem?

Feltehetően arról van szó, hogy a madárnak több fajtája létezik. És a madárfajták fajtáinak – például a varjaknak – talán nincsen több fajtája? A varjú fogalmának is nagyon különféle fajtái vannak

(dolmányos, vetési stb.). A fajták „lefelé” irányuló generikus hierarchiája végtelen. Figyeljük meg:

madár
 varjúféle
 varjú
 vetési varjú
 vastagsőrű vetési varjú
 vastagsőrű kisalföldi vetési varjú
 vastagsőrű kisalföldi rikácsoló vetési varjú
 vastagsőrű kisalföldi tegnapelőtt még rikácsoló vetési varjú
 vastagsőrű kisalföldi tegnapelőtt még rikácsoló de ma már kimúlt vetési varjú
 stb.

A kisebb betűs kifejezések ugyan már nem rendszertani, hanem köznyelvi fogalmakat neveznek meg, de ugyanúgy fogalmat, akár összes fölérendeltje. Mindegyik fogalom terjedelméhez végtelen számú előfordulás tartozik, és mindegyik fogalomnak sok fajfogalma lehet – azaz végtelen „alakban” fordulnak elő.

Az sem érv, hogy az adott tezaszuszban bizonyos hierarchiaszinttől lefelé az adott fajfogalmak fajtáit már nem veszik fel, így például a varjú fajtáit se, és ezért szerepel a varjú egyes számban. A madár deszkriptorcikkét látva a felhasználó amúgy is tudja, hogy a madárnak van további alárendelt fogalma, a varjú deszkriptorcikkét látva pedig látja, hogy a vajúnak viszont nincsen további alárendeltje. Nem kell ehhez a madarakat többes számban, a varjút pedig egyes számban szerepeltetni.

Ráadásul az eljárásból az következne, hogy ha netán mégis lesz alárendeltje egy addig egyes számban szereplő lexikai egységnek, akkor annak megnevezését többes számúra kellene majd módosítani. Ki tartja ezt észben azok közül, akiknek gondjuk a gyakorlatban a tezaszusz karbantartása?

A többes számú alakokat általában a tárgyakat, fizikai dolgokat, élőlényeket jelentő szavak esetében szokták használni. Ennek a felesleges használatnak az az oka, hogy nem világos a fogalmi (*tartalmi*) és a *terjedelmi* szint közötti különbség. Minden fogalom (*tartalmilag*) egyedi, egyetlen, melyet nyelvi szinten megnevezés azonosít. A köznyelvben ez az azonosítás tökéletlen (a szinonimák, a homonimák és poliszémák következtében), információkereső nyelvekben viszont egyértelmű, éppen ezért számítanak ezek a nyelvek szabványosítottoknak. Ezzel szemben a fogalmak *terjedelme* átfogó, azaz egy fogalom általában egynél több előfordulást, egyedet reprezentál (no-

ha lehet olyan fogalom – pl. kentaur –, amelynek terjedelme üres, azaz elvileg valami nem létezőt reprezentál). Tárgyak esetében ez elég világosan érzékelhető, folyamatok, tulajdonságok és állapotok, valamint elvont fogalmak esetében alig. Ez a tárgyakra vonatkozó „sok”, melyet a fogalom reprezentál, kelti azt a benyomást a tezauruszszervezőkben, hogy többes számban kell szerepeltetnie a tárgyakat reprezentáló fogalmakat.

A szavak csoportjai

A tezauruszok lexikai egységeit bizonyos szempontból valóban célszerű megkülönböztetni, mégpedig annak érdekében, hogy adott esetben a közneveket megfelelő, számítástechnikailag is kezelhető minősítéssel elválasszák a tulajdonnevektől, továbbá elválasszák a szaktárgyszavakat a formai tárgyszavaktól (dokumentumtípust, műfajt jelentő lexikai egységektől). Ez azért szükséges, mert a könyvtári rendszerekben a tulajdonneveket, a szaktárgyszavakat és a formai tárgyszavakat mind a bibliográfiai, mind a besorolási adatcsereformátumnak más-más mezőiben kezelik. A tezaurusz könyvtári rendszerbe való átvétele esetében tehát a fentiek szerinti csoportosításban kell tudni átadni az állományt (2. táblázat).

A tulajdonneveken belül külön kell kezelni a személy- és a testületi neveket, a rendezvényneveket

meg a címeiket a földrajzi tulajdonnevektől. Az első négy névtípus neveit általában nem foglalják tezauruszokba, csak a földrajzi tulajdonneveket: a konkrét természetföldrajzi képződmények, a működési-igazgatási, valamint a településneveket (pl. Mátra, Pest vármegye, Esztergom). A földrajzi tulajdonnevek nem cserélendő össze az általános földrajzi nevekkel (pl. hegység, vármegye, város).

Előfordul ugyan személynév vagy testületi név tezaurusz, de az ilyen tezauruszok neveket jelentő lexikai egységeit tartalmi feltárás céljára nem ugyanazokban az adatmezőkben kezelik a könyvtári rendszerek, mint a szaktárgyszavakat és a formai tárgyszavakat, illetve a további földrajzi neveket. Ezért a személynév- és a tulajdonnév-tezauruszok alapvetően elkülönülnek az információkereső tezauruszok többi válfajától.

A személynéveket és a testületi neveket ugyanis, ha velük a dokumentum tárgyait, témáit osztályozzák, a bibliográfiai szabványok szerint tárgyi melléktételként kell fölvenni (az adatcsere-formátumok 610-es adatmezőjében). A földrajzi nevek tárgyi melléktételének az adatcsere-formátumok egy másik adatmezője, a 651-es felel meg; a szaktárgyszavak tárgyi melléktételének a 650-es, a formai tárgyszavakénak pedig a 655-ös adatmező.

2. táblázat

Az információkereső-nyelvi szavak típusai és adatmezői a bibliográfiai rekordban

Az információkereső-nyelvi szavak mezői a bibliográfiai rekord adatcsere-formátumában			
<i>hívó-jel</i>	<i>megnevezés</i>	<i>bibliográfiai típusa</i>	<i>példa (megjelenített forma)</i>
600	személynév	tárgyi melléktétel	Katona József (1791–1839)
610	testületi név	tárgyi melléktétel	Petőfi Sándor általános iskola (Zalaegerszeg)
611	rendezvénynév	tárgyi melléktétel	Arany János centenáriumi emlékülés (1982) (Nagykőrös)
630	egységesített cím	tárgyi melléktétel	Biblia (Újszövetség)
650	tárgyszó	tárgyi kifejezés	gyalogtúra, építészet, véres csütörtök (Szentpétervár) (1905)
651	földrajzi név	tárgyi melléktétel	Bázel, János-hegy (Budapest), Szeged. Tisza Lajos körút
653	szabad tárgyszó	mutatószó	rézcső, inzulinfüggést eredményező terápia, sárgányfog,
655	formai tárgyszó	tárgyi melléktétel	regény, zenei etűd, kutatási jelentés, képeslap, mozijegy
656	foglalkozás tárgyszó	tárgyi melléktétel	erdészeti munkás, asztalos, könyvelő, egyetemi tanár
657	dokumentum célja, rendeltetése	tárgyi melléktétel	utcai szociális gondozás, baleset-elhárítás, éves leltár
658	oktatási (tananyag) tárgyszó	tárgyi melléktétel	mérnökképzés, hegyi mentés, készségfejlesztés, öntészet
662	hierarchikus földrajzi név	tárgyi melléktétel	Egyesült Államok, Kalifornia, Los Angeles (county), Los Angeles, Little Tokyo

Az adatsere-formátumban további – ritkán alkalmazott – tárgyszótipusok is szerepelnek (foglalkozási tárgyszó, dokumentum célja/rendeltetése tárgyszó stb.).

A személyneveket, testületi neveket, rendezvényneveket, címeket, földrajzi tulajdonneveket, szak-
tárgyszavakat, formai tárgyszavakat és egyéb tárgyszótipusokat összefoglalóan információkereső-nyelvi szavaknak nevezzük.

Ha bármilyen tezauszus gyakorlati alkalmazására kerül sor, a tezauszus eredményes átvételéhez a fentiek szerinti szétválasztás elengedhetetlen. A hallgatók gyakorlati felkészítését szolgálja, hogy a diplomatezauszusokban erre is gondolniuk kell. A 2. táblázatban a bibliográfiai adatsere-formátum

3. táblázat

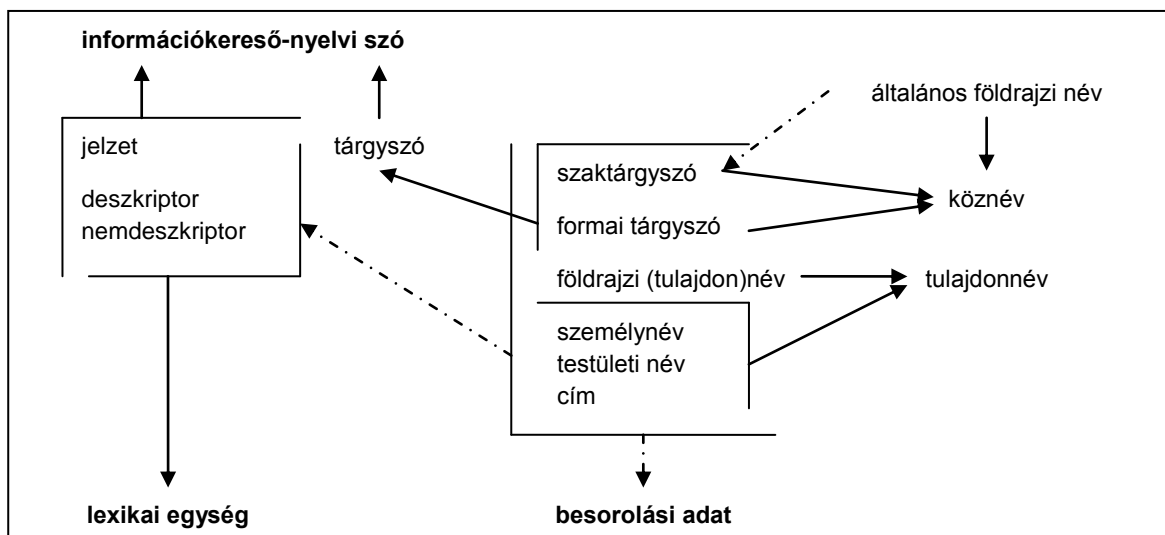
Az információkereső-nyelvi szavak típusai és adatmezői a besorolási rekordban. A bibliográfiai rekord 656–662-es mezőinek nincsenek megfelelői

Az információkereső-nyelvi szavak mezői a besorolási rekord adatsere-formátumában			
<i>hívójel</i>	<i>megnevezés</i>	<i>bibliográfiai típusa</i>	<i>példa (megjelenített forma)</i>
100	személynév	tárgyi melléktétel	Katona József (1791–1839)
110	testületi név	tárgyi melléktétel	Petőfi Sándor általános iskola (Zalaegerszeg)
111	rendezvénynév	tárgyi melléktétel	Arany János centenáriumi emlékülés (1982) (Nagykőrös)
130	egységesített cím	tárgyi melléktétel	Biblia (Újszövetség)
148	kronologikus tárgyszó	tárgyi melléktétel	19. század, véres csütörtök (Szentpétervár) (1905), máj. 1.
150	tárgyszó	tárgyi kifejezés	gyalogtúra, építészet, asztalos, baleset, mérnökképzés
151	földrajzi név	tárgyi melléktétel	Bázel, János-hegy (Budapest), Szeged. Tisza Lajos körút, Little Tokyo
155	formai tárgyszó	tárgyi melléktétel	regény, zenei etűd, kutatási jelentés, képeslap, mozijegy

szerint kezelt információkereső-nyelvi szavak típusai láthatók.

A 3. táblázatban az előbbi típusoknak a besorolási adatsere-formátum szerint kezelt adatmezői láthatók. A formátumok sajátja, hogy a besorolási formátumban kevesebb típust különböztetnek meg, mint a bibliográfiaiban. Ennek elemzése önmagában is megér egy szakdolgozatot vagy diplomamunkát és azt követő publikációt.

A 7. ábrában címkézett, irányított gráf formájában is összefoglaltuk a fenti táblázatokban szereplő információkereső-nyelvi szavakat. A gráf egyben példa a tezauszus grafikus részére is. Diplomamunkában a szerkesztés támogatására hasznos, ha a nehezebben strukturálható szócsoportok esetében grafikus formában tekintik át a relációkat.



7. ábra Az információkereső-nyelvi szavak típusai címkézett irányított gráfban

A folytonos egyirányú élek a generikus, a pontvonalas élek a „váltak valamivé/felhasználható valamire” E/R oksági tezauruszrelációt jelölik. A félnégyzetekkel összefogott lexikai egységek mindegyikére érvényes a kapcsolódó reláció. Például a szaktárgyszó, a dokumentumtípus, a földrajzi tulajdonnév, a személynév, a testületi név, a cím válhat besorolási adattá, illetve deszkriptorrá vagy nemdeszkriptorrá.

A generikus reláció ismerete

Az angol relációjeleket használó tezauruszok egy részében általában csak az átfogóbb–szűkebb (broader term, narrower term) és a rokonsági relációt (related term) szokták használni a deszkriptorok között. Ez azonban elégtelen a lényeges összefüggések elkülöníthető rögzítésére.

Az elnagyolt relációszerkezet következménye, hogy ún. „laza” hierarchia keletkezik. Akárcsak az ETO-ban, az ilyen tezauruszokban is olyan fogalmakat rendelnek egymás alá, amelyek teljesen más fogalmi kategóriákba tartoznak (fogalmi kategória például az „anyag [körvonalazatlan anyagot reprezentáló fogalom, mint „homok”, „tej”], a tárgy [körvonalazott dolgot reprezentáló fogalom, mint „kő”, „kuvasz”, „gép”], a folyamat [dinamikus változást reprezentáló fogalom, mint „esés”, „robbanás”, „járás” stb.]).

Az egyik tezauruszban például a „máz” a „szilikátiparnak” van alárendelve, holott a máz nem fajtája semmiféle iparnak.

A generikus reláció kezelésében azonban akkor is bizonytalanság mutatkozik, ha történetesen a szabvány erre meghatározott speciális relációtípusát alkalmazzák.

Az egyik tezauruszban például az összes *szerszámot* partitív relációban kapcsolták össze a *szerszám* fajtáival, illetve az általánosabb fogalmával (a generikus fölérendeltjével), holott ezt a generikus relációval (F/A) kellett volna megvalósítani (8. ábra). A *kés* például nem része a *konyhaeszköznek* vagy a *vágószerszámnak*, hanem a fajtája. A fajta nem azonos a résszel! A *nyél* ezzel szemben valóban a része a *késnek*. A 8. ábra első deszkriptorcikkében a többes szám felesleges és ellentmondásokhoz vezető alkalmazása is látható: az egyik fölérendelt többes számban szerepel (konyhai eszközök), a másik viszont nem (vágó-

szerszám). Ugyan miért? A vágószerszámnak is van tucatnyi fajtája, nem csak a konyhaeszköznek.

Hiba sok tezauruszban, hogy a lexikai egységek mindig nagybetűvel kezdődnek. A példaként választott tezaurusz kifejezései nem tulajdonnevek, hanem fogalmat neveznek meg (ezek a nevek általános nevek, illetve köznevek). A közneveket a magyar helyesírási szabályok szerint kis kezdőbetűvel írjuk.

hibás:		helyes:	
4230			
KÉS		kés	
H	Messer	H	Messer
T	Vágószerszám	F	vágószerszám
	Konyhai eszközök		konyhaeszköz
P	Nyél	P	nyél
E	Acél	E	acél
	Fa		fa
	Műanyagok		műanyag
	Vágás		vágás
	Vágási technika	X	vágási technika

8. ábra Hibás és helyes deszkriptorcikk. A bal oldali deszkriptorcikk élén, a vezérszó fölött látható a lexikai egység teljesen felesleges azonosítója

A generikus relációk elnagyolt kezelése megtévesztő. Nem teszi lehetővé, hogy adott esetben kihasználhassák automatizált formában (pl. egy ontológiában, de akár egy korszerűbb keresőrendszerben is) a generikus relációban szereplő logikai lehetőségeket, például a fölérendelt automatikus bevonását a keresésbe (generic posting, upposting). A tulajdonságöröklődés, mely mind az online adatbázisokban, mind az ontológiákban a keresés egyik legfontosabb alapja, kizárólag a generikus reláció esetében áll fenn. Ha valaki a „máz” szóval keresve nem talál elég dokumentumot, a „bevonat” kifejezéssel keresve releváns találatokhoz juthat. Ha viszont a „máz” helyett a „szilikátipar” kifejezéssel keres, aligha jut használható dokumentumokhoz.

Ha a szóban forgó tezauruszban a szabványos relációjeleket alkalmazták volna (lett legyen magyar vagy angol¹⁰), ez a probléma talán nem alakult volna ki. A máz ugyan a szilikátipar (meg a festészet, meg a címertan stb.) szakterületén belül is előforduló fogalom, de ezekben az esetekben az összesség–elemek reláció áll fenn, melyet magyarul az „egésze–része” (magyar szabvány szerint: T, illetve P), angolul a „broader term partitív”, illetve „narrower term partitív” szabványos angol relációnévvel adhatunk meg (angol jele BTP, illetve NTP).

A magyar és az angol szabvány teljes relációválasztéka a 4. táblázatban látható. Megadtuk a tezaurusokra (is) vonatkozó HUNMARC adatcsereformátum [2] almezőjeit is, hogy látható legyen: az adatcsere-formátumok létezése ma már megkerülhetetlen a tezaurusok világában. A besorolási adatcsere-formátumban a táblázatban szereplő relációknál több relációtípus szerepel, az alábbiak az ismertebbek Magyarországon.

A szabvány egyébként további speciális reláció felvételét is lehetővé teszi. Megnézhető ez többek között a Köztauruszban, melyet az OSZK honlapjáról bárki letölthet, és elérhető az állománya a *Magyar Elektronikus Könyvtáron* keresztül is [5]. A Köztauruszban például megtalálható a lexikai egységek és a vonatkozó ETO-jelzet közötti összefüggés is, melyet külön reláció jelöl (9. ábra).

4. táblázat

A magyar [6] és az angol [3] [4] tezaurusrelációk összehasonlító táblázata

Relációtípus	HUNMARC almezője	magyar relációnév	jele	angol relációnév	jele
deszkriptora	450##a	lásd	L	see / use	see / use
nemdeszkriptora	450##a	helyett (lásd innen)	H	used for	used for, UF
deszkriptora más deszkriptorral együtt	450##ws	lásd ÉS	L&	see AND / use AND	see AND / use AND
nemdeszkriptor más deszkriptorral együtt	450##wt	helyett ÉS (lásd innen ÉS)	H&	used for AND	used for AND
deszkriptora vagyilagosan	450##wu	lásd VAGY	L	see OR / use OR	see OR / use OR
nemdeszkriptora vagyilagosan	450##wv	helyett VAGY (lásd innen VAGY)	H	used for OR	used for OR
generikus reláció	450##wg	átfogóbban/általánosabban	F	broader term generic	BT
	450##wh	speciálisabb/fajtája	A	narrower term generic	NT
partitív reláció	450##wj	egésze/teljes	T	broader term partitív	BTP
	450##wk	része/eleme	E	narrower term partitív	NTP
oksági reláció	450##wp	eredménye/folytatása	R	broader term causal	BTC
	450##wr	kiindulása/előzménye	E	narrower term causal	NTC
egyéb asszociatív reláció	450##wm	egyéb rokonsága/lásd még	X	related term	RT

címer	
M:	<i>Hagyományosan pajzs formájú, valamely közösség megkülönböztetésére használt, alakos, színes jelvény</i>
CF	342.22 676.818
H	címerpajzs
F	állami szimbólum emléma
A	magyar címer
T	címertan
X	családfa falerisztika muzeológia pecsét zászló
≠	agancs cégtábla

címer	
M:	<i>Bizonyos szabályok szerint, meghatározott salakokból alkotott képek, melyeket egyes családok vagy testületek: államok, megyék, városok, mint saját és állandó jelvényeiket, régi jog vagy fejedelmi adományozás alapján, törvényesen használnak. Ismertető jelei: hogy az azt alkotó kép pajzsba legyen foglalva, és az ember minden jogaival együtt átörökölheto legyen</i>
BTP	címertan
NT	államcímer beszélő címer céhcímer községi címer megyei címer polgári címer városi címer teljes címer
RT	családjegy kereskedőjegy pecsét

9. ábra A címer lexikai egység deszkriptorcikke az egyetemes Köztauruszban (bal oldalt) és egy diplomatezauruszként készült speciális címertani tezauruszban (jobb oldalt). A megjelenítéshez a Relex tezauruszkezelőt használtuk

A Köztauruszban a címerhez több általánosabb, nem a címertannal összefüggő fogalom kapcsolódik, ami indokolt, hiszen egyetemes, átfogó tezauszról van szó. A címer meghatározásának forrása a *Magyar Értelmező Kéziszótár*.

A Köztaurusz CF (csúcsfogalma) relációja kapcsolja össze a deskriptort mint ETO-mutatószt azokkal a jelzetekkel, melyek a címer fogalmára vonatkoznak. A ≠ relációjellel kapcsolódnak a Köztauruszban azok a lexikai egységek, amelyeknek más az értelme (az agancsra és a cégtáblára is mondják, hogy „címer”, azaz a „címer” szó homonima).

A címertani tezauszban indokoltan több speciális fogalom kapcsolódik a címerhez. A címer címertani meghatározása általában az információkereső tezauszok számára túl hosszú, inkább terminológiai vagy makrotezauszba való. Információkereséskor, online körülmények között ilyen hosszú meghatározás nehezen olvasható.

Ha nem használják a partitív egésze relációt (az angolban BTP), a címert hibásan a BT relációval (főlérendeltje) kellett volna összekapcsolni a címertannal, holott a címer nem címertan, hanem a címertan szakterületének egyik eleme, tárgya. A tezauszokban az ilyen összefüggést (ismeretterület és tárgya, eleme) egységesen a partitív relációval fejezik ki.

A megjegyzések

A lexikai egységekhez több, különböző típusú megjegyzés, magyarázat kapcsolható. A fontosabb típusokat a HUNMARC besorolási adatok adatsere-formátuma is tartalmazza [2] (5. táblázat). Ezeknek természetesen megvannak a MARC21 szerinti angol megfelelői is, az olvasó utánanézhethet a formátum honlapján.

A diplomatezauruszok készítői csak ritkán élnek az alkalmal, hogy minden lexikai egységhez meg-

jegyzést fűzzenek, holott az általuk készített mikrotezauruszok nagyon speciális fogalmai különösen igénylik ezt. Ha használnak is, ez kizárólag egyetlen típusra, az általános megjegyzésre szorítkozik, melyet meghatározás céljára vesznek igénybe.

Mivel a tezauszok nem terminológiai szótárak, hanem elsősorban információkereső célra készülnek (noha lehetségesek terminológiai célú tezauszok is), ezért bennük a meghatározások legrövidebb, gyorsan áttekinthető formáját, a generikus meghatározást célszerű használni. Ennek lényege, hogy a definíció mondata mindig tartalmazza a meghatározott fogalom generikus főlérendeltjét. Formája: „A [van] B”. Kerülni kell a terjengős, többnyire „[A] olyan B, amely ...” formát, ami a diplomatezauruszokban sajnos a leggyakoribb eset.

Példák helytelen meghatározásokra:

machete

M: *Egy rövid nyelvű, erős pengéjű eszköz, cukornád-aratók használják Közép- és Dél-Amerikában*

közigazgatási bíróság

M: *Olyan testület, amely rendes bíróságot szervezetétől eltérő szervezetben az állampolgároknak a közigazgatással, az önkormányzatoknak a kormányval szembeni ügyeiben dönt; döntéseivel szemben nincs további fórum fellebbezni*

Helyesen:

machete

M: *Közép- és dél-amerikai cukornádaratók rövid nyelvű, erős pengéjű szerszáma*

közigazgatási bíróság

M: *A közigazgatási döntések ellenőrzésére hivatott, a rendes bíróságok szervezetétől eltérő szervezetű, az állampolgároknak a közigazgatással, az önkormányzatoknak a kormányval szembeni ügyeiben mint utolsó fórumként döntő különös bíróság*

5. táblázat

A tezausz megjegyzései fontossági sorrendben

Hívójel és indikátor	Almező	Tartalma	Rövid jele a tezauszcikkb ¹¹	Fehasználóbarát neve a tezauszcikkb ¹¹
680 ##	\$i	Általános megjegyzés, magyarázat	M:	Megjegyzés:
690 ##	\$a	Használati megjegyzés	H:	Használat:
678 ##	\$a	Történeti megjegyzés	T:	Történet:
688 ##	\$a	Megjegyzés az alkalmazás történetéről	Vált.:	Változás:
670 ##	\$a	Forrás	Forrás:	Forrás:
667 ##	\$a	Belső megjegyzés	B:	Belső megjegyzés:

Ha elkerülhetetlen, hogy a meghatározás hosszú legyen, a generikus definíciót képviselő mondat kiegészíthető további mondatokkal. Például:

beálló cselekvésű főnévi igenév

M: *A létigéből egyes nyelvjárási területeken kifejlődött, összetett igenévi forma, mely a létige jövő idejű alakjának -ni képzős származékából és ragozott köznyelvi, jövő idejű alakjából áll. Ez utóbbi fejezi ki a történés idejét (lenni lesz)*

beálló cselekvésű főnévi igenév

M: *A létigéből egyes nyelvjárási területeken kifejlődött, összetett, a létige jövő idejű alakjának -ni képzős származékából és ragozott köznyelvi, jövő idejű alakjából álló igenév. Ez utóbbi fejezi ki a történés idejét (lenni lesz)*

A besorolási adatcsere-formátum és a diplomatezauruszok

A formátum adta lehetőségek az oktatásban

Aki tezaurszt szerkeszt, annak tudnia kell, hogy a terméket az adatcsere-formátumok számára alkalmas szerkezetben kell elkészítenie ahhoz, hogy a könyvtári rendszerekben használni lehessen. A formátumokhoz való hozzáértés ma könyvtári szakterületen lényegében piacképes tudás.

Fontos, hogy az egyetemi oktatás keretében elkészülő tezaurszokban és velük összefüggő dolgozatokban is megjelenjen a besorolási adatcsere-formátum, mint a tezaursz belső, formális kezelésre alkalmas szerkezete. A hallgatók e formátumon keresztül tapasztalhatják meg a tezaurszhasználat gyakorlatát. Bárki, aki könyvtári munkahelyén bibliográfiai, osztályozási, de akár gyarapítási munkakört lát el, munkába lépésének első pillanatától kezdve azokkal a képernyőfelületekkel szembesül, amelyek a bibliográfiai és a besorolási adatcsere-formátumokon alapulnak. A bibliográfiai és a tartalmi feldolgozás minden adata a formátumok előírta mezőkbe rögzítendő.

A besorolási adatcsere-formátum valójában a tezaurszokra vonatkozó tudás tárháza is. Aki ezt a formátumot ismeri, mindent tud a tezaurszok és a tárgyszórendszerek szerkezetéről. Mégis kevés az olyan könyvtári szakember, aki maradéktalanul eligazodik bennük. Ebből következik, hogy számos, egyébként a formátumok adta minden lehetőségre felkészített könyvtári rendszert hiányosan adaptálnak, mert maguk a megrendelő, alkalmazó

könyvtárosok nem ismernek minden, a formátumokban szereplő, használható adatmezőt és relációt. Ezért a tezaurszokban rejlő lehetőségek is kihasználhatlanok maradnak. Mind a professzionális, mind a végfelhasználó számára ma kiáltoan hiányzanak az online mutatókból a magyarázatok, a gazdagabb utalások és hivatkozások, melyek mind létrehozhatók lennének a formátum alapján. Ennek az osztályozás minősége is kárát látja, mert a szakemberek a napi munkájukban nem kapnak elég támogatást a szegényes online mutatóktól.

Ez nemcsak a tezaurszokra és a tárgyszórendszerekre vonatkozik, hanem az Egyetemes Tizedes Osztályozásra és annak jelzeteire is. Az ETO jelzetrekordjai is besorolási rekordok, rájuk az osztályozási jelzetek adatcsere-formátuma vonatkozik (MARC Classification Format) [2]. Ha a jelzetrekordokat e formátum szerint töltik be a könyvtári kezelőrendszerbe, ugrásszerűen megnő az ETO használhatósága, mert a felhasználó ettől kezdve nemcsak a jelzetet látja az online mutatóban, hanem a jelzet magyarázatát is, és ráadásul e jelzetmagyarázatok szövegének szavai szerint is kereshet.

Az adatcsere-formátumok számos egyetemen a hallgatók által készülő munka tárgyai lehetnének.

Ma már minden valamire való nemzeti könyvtárnak van saját besorolási adatállománya. A *Kongresszusi Könyvtár* (MARC21), a *Spanyol Nemzeti Könyvtár* (IBERMARC) és a *Francia Nemzeti Könyvtár* (GALLICA) besorolási adatállományai, a *Finn Nemzeti Könyvtár* (FINMARC) tezaurszai mintaszerűek [20]. (A MARC21 – ha nem is minden részletében, de – le van fordítva német nyelvre is [2].)

Diplomamunkák és szakdolgozatok kimeríthetetlen tematikus forrásai ezek az állományok. Számos ötletet adhatnak a motiváltabb hallgatói munkához.

Év közbeni munkák lehetnek például:

- különböző típusú kifejezések (személynevek, testületi nevek, rendezvénynevek, címek, földrajzi tulajdonnevek, szaktárgyszavak, formai tárgyszavak stb. tezaurszcikkeknek, és e tezaurszcikkeknek megfelelő besorolási rekordok mezős szerkezetének elkészítése;
- ETO-jelzetrekord és e jelzetrekordnak megfelelő MARC21 osztályozási besorolási rekord mezős szerkezetének elkészítése;

- a besorolási rekord 008-as adatmezőjében szereplő egyes pozíciók értékelése, és példák megadása az egyes pozíciókra;

Év végi vagy szakdolgozati munkák lehetnek például:

- a fenti nemzeti besorolási adatállományok és az *Országos Széchényi Könyvtár* magyar nyelvű besorolási adat-állományának összevetése tanulságokkal szolgálhat hallgatónak, tanárnak és szakembernek egyaránt;
- elkészíthetők kétnyelvű teauruszrészletek mind adatcsere-formátumok, mind szabványosan megjelenített teauruszcikkek formájában;
- megvizsgálhatók ETO-jelzetek és besorolási adatok közötti összefüggések;
- összehasonlíthatók a Köztaurusz [5] és a nemzeti könyvtárak teauruszai;
- külön tanulmányok tárgyai lehetnek a besorolási adatcsere-formátumok összetett magyarázatos utalói (a 260-as, a 360-as és a 66X mezők), melyek hiányoznak a teauruszszabványokból, holott szükség lenne rájuk; ezeknek az utalóknak a használata a Kongresszusi Könyvtárban tanulmányok sorát indokolná;
- fontos tanulmányok születhetnének a *Kongresszusi Könyvtár* és a debreceni *Kossuth Lajos Tudományegyetem Könyvtárának*, meg a *Szegedi Egyetem Könyvtárának* összetett tárgyszavas gyakorlatáról, összevetve például a Köztaurusz lexikai egységeinek szerkezetével (az összetett tárgyszavak alkalmazásának a szaknyelven Magyarországon jóformán elsorvadt)¹².

A végzős hallgatók ösztönzése az ilyen munkákra nemcsak azzal jár, hogy felkészült szakemberek lesznek, hanem a magyar könyvtártudományt is gazdagítják megtermékenyítő felismerésekkel. Hiszen meglehetősen ismeretlen kérdésekről van szó Magyarországon.

A teauruszcikk a formátum szerint

A 3. és 4. ábrában már bemutatottuk, hogyan fest egy olyan diplomatezaurusz besorolási tétele a HUN-MARC adatcsere-formátumban, amelyben megoldják, hogy felhasználóbarát relációjelekkel is megjelenhessen adott kezelőrendszerben a teaurusz.

A következőkben (10. és 11. ábra) bemutatunk egy, az OSZK-teauruszból származó, és egy jogi teauruszba is átvett teauruszcikket, továbbá egy ETO-jelzetrekordot besorolási adatcsere-formá-

tumban, és látható a megjelenítésük is. A példákat úgy választottuk, hogy a teauruszcikk deszkriptora egyben az ETO-jelzet mutatószava is, azaz a deszkriptor és a jelzet a formátumokon keresztül összekapcsolódik.

A példákat a Kongresszusi Könyvtár és a MARC21 gyakorlatán keresztül meghonosodott formában mutatjuk be. Az adatmezőt háromjegyű szám, az almezőt egy karakter azonosítja (hogy az adott karakter az almezőhívójel szerepét játssza, azt szövegben az előtte álló \$ jel jelzi). A # indikátorok esetében betöltetlen pozíciót jelent. A 008-as mezőben a félkövér # jelek a HUNMARC szerinti betöltetlen (felhasználatlan) pozíciókat jelölik, a dőlt # jelek pedig minősítés képviselő üres helyek.

A teauruszok deszkriptorcikkeinek (a formátum nyelvén: kitüntetett tételfejű tételeinek) jellegzetesége, hogy több 4XX és 5XX utaló mezőt tartalmaznak. A deszkriptor (a tételfej) magyarázata (680), használati előírása (690), története (678) megjegyzésekben szerepel, az utalásokat (kapcsolatokat) ugyancsak megjegyzésekben indokolják (667), ha kell. Gyakorik a forrásadatok is (670-es és 675-ös mező).

A besorolási tétel 008-as mezőjének 11. pozícióján mindig szerepel a vonatkozó tárgyszórendszer/teaurusz jelzete (a példánkban o=az OSZK-teaurusz jelzete), a kitüntetett tételjelek mindig használhatók tárgyi melléktételben (008/15=a), nem kormányzati jellegűek (008/28=#), többnyire van ellenőrzött utalójuk a 4XX mezőben (008/29=a), deszkriptor–deszkriptor kapcsolatuk az 5XX mezőkben (általában \$w vezérlő almezővel megadva), és az elfogadottság mértékére vonatkozóan fontos adat, hogy mi a forrása (a példánkban 008/39=#, azaz nemzeti könyvtári).

A teauruszok deszkriptorcikkeinek többnyire viszonylag sok (8–10) kapcsolata van más deszkriptorokkal (550-es mező). A viszonylag sok kapcsolat miatt a \$a almező által megadható „lásd még” kapcsolatok áttekinthetetlenül összefolynának. Ezért alkalmazzák a \$w vezérlő almezőt a specifikus kapcsolattípusok kifejezésére (generikus, partitív, oksági, egyéb rokonsági), hogy a kapcsolatok a deszkriptorcikk vezérszava alatt ne a kapcsolódó kifejezések pusztá betűrendjében, hanem a relációtípusok szerinti csoportokban, jól áttekinthetően jelenjenek meg. A kapcsolattípusok sorrendje a megjelenítésben a generikus relációval (\$w g és h) kezdődik, mivel ez a legfontosabb

információkereső-nyelvi reláció, és a szemantikai fontosság sorrendjében következnek egymás után.

A 750-es tételfejkapcsoló mezőn keresztül a tételfej az ETO osztályozási adatsere-formátumában szereplő ETO-jelzet tételfejhez kapcsolódik (ez a

kapcsolat két külön besorolási rendszer közötti összefüggés). A \$2-es almezőben az ETO forrásjele (eto) szerepel. A fordított irányú kapcsolatot az ETO-jelzettel mint kitüntetett tételfejvel, és magát az ETO-jelzetrekordot az „Osztályozási adatok adatsere-formátuma” szabályozza.

Tag ii \$nz n # 5.
000 #####nz 22#####n 4500
001 <rekordazonosító>
005 20080713233723.0
008 080603 #n#an z##ba #n### ##### ###a# ana## ####d
040 ## \$aMBA \$bhu \$fJogi és igazgatásügyi tezaurusz
150 ## \$aközigazgatási bíróság
450 ## \$wy \$államigazgatási jogszolgáltatás
450 ## \$wy \$aközigazgatási bíraskodás
450 ## \$wy \$aközigazgatási kollégium
550 ## \$wg \$abíróság
550 ## \$wk \$aközigazgatási jog
550 ## \$wk \$abírósági felügyelet
550 ## \$wk \$aközigazgatási ellenőrzés
550 ## \$wg \$aválasztói kifogás
550 ## \$wm \$aközigazgatási peres eljárás
670 ## \$aAz 1991. évi XXVI. törvény a közigazgatási határozatok bírósági felülvizsgálatának kiterjesztéséről.
670 ## \$almre M.: Közig. Bíraskodás, 2008
670 ## \$aOSZK-tezaurusz deskriptora, átvéve 2008 \$cUR
678 ## \$aMagyarországon 1883-1949, majd 1990-től fennálló bíróság
680 ## \$aA közigazgatási döntések ellenőrzésére hivatott, a rendes bíróságok szervezetétől eltérő szervezetű, az állampolgárnak a közigazgatással, az önkormányzatoknak a kormánnyal szembeni ügyeiben mint utolsó fórumként döntő különös bíróság
690 ## \$aAz 1949-1990 közötti időszak államigazgatási határozatokkal szembeni eljárások, felülvizsgálatok esetén a „jogi felülvizsgálat”, vagy a „jogi eljárás” és az adott kérdés deskriptora („lakásügy”, „illetékügy” stb.), továbbá adott esetben az „egyeztető bizottság” használandó
750 ## \$a342.565.4 \$2eto
750 ## \$a347.998.95 \$2eto
750 ## \$a351.95 \$2eto
750 ## \$administrative court \$0(euvoc)01539 \$2euvoc
közigazgatási bíróság
M: A közigazgatási döntések ellenőrzésére hivatott, a rendes bíróságok szervezetétől eltérő szervezetű, az állampolgárnak a közigazgatással, az önkormányzatoknak a kormánnyal szembeni ügyeiben mint utolsó fórumként döntő különös bíróság
H: Az 1949-1990 közötti időszak államigazgatási határozatokkal szembeni eljárások, felülvizsgálatok esetén a „jogi felülvizsgálat”, vagy a „jogi eljárás” és az adott kérdés deskriptora („lakásügy”, „illetékügy” stb.), továbbá adott esetben az „egyeztető bizottság” használandó
Tört: Magyarország 1883-1949, majd 1990-től fennálló bíróság
Forrás: Az 1991. évi XXVI. törvény a közigazgatási határozatok bírósági felülvizsgálatának kiterjesztéséről. Imre M.: Közig. Bíraskodás, 2008 OSZK-tezaurusz deskriptora, átvéve 2008 (MÉ)
H államigazgatási jogszolgáltatás közigazgatási bíraskodás közigazgatási kollégium
F bíróság
T közigazgatási jog
R bírósági felügyelet közigazgatási ellenőrzés választói kifogás
ETO 342.565.4. 347.998.85 351.95

10. ábra A deskriptor besorolási rekordja a HUNMARC adatsere-formátuma szerint, és megjelenített deskriptorcikke

A 750-es mezőben lett megadva a többnyelvű EUROVOC jogi tezaurusz megfelelő deskriptorával is a kapcsolat. A \$2-es almezőben az EUROVOC forrásjele (euvoc) szerepel. Ezt a kapcsolatot a 10. ábra példájában nem jelenítettük meg. Az ilyen megjelenítések mindig rendszertől függenek, illetve opcionálisak adott rendszerben; nem minden mai (egyébként a MARC21 igényeit kielégíteni képes) könyvtári rendszert adaptáltak úgy (általában tájékozatlanság következtében) a magyarországi könyvtárakban, hogy vele például a 750-es mezővel, de akár az 550 \$w, és a 6XX megjegyzésmezőkkel kapcsolatos fontosabb lehetőségek kihasználhatók lennének.

A besorolási adatcsere-formátumok és a tezauruszkészítés oktatását a MARC21 színvonalán megvalósító könyvtárosképzés valójában minimumfeltétel.

Az ETO-jelzetrekord a 2000. évi kiadásból való. A 153-as mező \$a almezőjében látható a jelzet, a \$j almezőjében pedig a jelzet magyarázata, ahogy az az ETO táblázatában szerepel. Az 553-as mezők „lásd még” kapcsolatban álló jelzeteket tartalmaznak, melyeket az ETO táblázataiban a → relációjel jelöl. A 750-es mezőben szerepel a tezaurusz mutatószóként kapcsolódó deskriptora, a 753-as

mezőkben pedig az ETO betűrendes mutatójának mutatószavai.

Az a diplomatezaurusz-készítő hallgató, aki néhány tezauruszcikkének elkészíti a besorolási adatrekordját (és azt tezauruszának előszavában bemutatja), joggal tekintheti magát a kor szakmai színvonalán álló könyvtári szakembernek.

Epilógus

Ha én diplomatervként tezauruszt készíthetnék, amellett, hogy a Relex tezauruszszerkesztőt [14] használnám, felhasználnám a Köztauruszt is. Nem azért, mert ez az egyetemes tezaurusz annyira tökéletes, netán speciális kifejezéseket is tartalmazna, hanem azért, mert praktikus dolog kapcsolódni olyan tezauruszhoz, amelyet Magyarországon már sok könyvtárban használnak.

Olyan tezauruszt készítenék, amelynek csúcskifejezései egyben a Köztaurusz [5] lexikai egységei is lehessenek, továbbá úgy állapítanám meg, hogy mi legyen deskriptor és mi legyen nemdeskriptor, hogy az összhangban legyen a Köztaurusz nemdeskriptoraival és deskriptoraival (azaz a redukált fogalmi szerkezetével).

Tag ii \$nz n # 5.	
000 *****nz 22*****n 4500	
001 <rekordazonosító>	
005 20080713233604.0	
008 000103 aaaaa aaa	
084 8# \$audc \$cE&C 15 2000	
153 ## \$a351.95 \$jÁllamigazgatási jogszolgáltatás. A hatalmi ágak konfliktusa. Közigazgatási államtanács	
553 0# \$a342.52 \$jA törvényhozó hatalom	
553 0# \$a342.52 \$jKözigazgatási bíróság	
683 0# \$aA volt szocialista országok esetén célszerű az időt/időszakot is osztályozni	
750 #7 \$aközigazgatási bíróság \$2mba	
753 ## \$ahatalmi ágak konfliktusai (közig.)	
753 ## \$aállamigazgatási jogszolgáltatás	
351.95	
Államigazgatási jogszolgáltatás. A hatalmi ágak konfliktusa. Közigazgatási államtanács	
Használat:	A volt szocialista országok esetén célszerű az időt/időszakot is osztályozni
→ 342.52	A törvényhozó hatalom
→ 342.565.4	Közigazgatási bíróság
mutatószava	államigazgatási jogszolgáltatás
	hatalmi ágak konfliktusai (közig.)
tezaurusz	közigazgatási bíróság
Kiadás	E&C 15 2000

11. ábra Az előbbi deskriptorhoz kapcsolódó ETO-jelzet besorolási rekordja a MARC osztályozási adatcsere-formátuma szerint, és megjelenített formája

Ezáltal elérem, hogy a tezauszom felülről kompatibilis lenne egy mások által sokat használt tezausszal. Ez azért jó, mert szép is, nem csak hasznos.

Például egy ütőhangszerekről készült diplomatezauszban a húros ütőhangszerekre a chordofon, a rugalmas test hangadásával működő ütőhangszerekre az idiofon, és a kifeszített hátrtyák hangadásával működő ütőhangszerekre a membranofon megnevezést használták. Ez nem probléma mindaddig, amíg kizárólag ütőhangszerekről van szó. A felsorolt három típus azonban minden hangszerre érvényes, azaz a doromb is chordofon, noha nem ütőhangszer, a hegedű is idiofon, és létezik olyan membranofon, amely nem ütőhangszer. A Köztauruszban ennek megfelelően szerepelnek a kifejezések (azaz van például idiofon, és van idiofon ütőhangszer). Ezért annak érdekében, hogy az ütőhangszer-tezausz felülről kompatibilis legyen a Köztauruszsal, a három csúcsfogalmát kiegészítették az „...ütőhangszer” névrésszel: „chordofon ütőhangszer”, „idiofon ütőhangszer” és „membranofon ütőhangszer”. Ez a három kifejezés ugyanis alárendeltje lehet a Köztaurusznak; zenei gyűjtőkörben tehát az ütőhangszer-tezausz akadálytalanul beilleszthető az átfogó Köztauruszba.

Jegyzetek

- 1 Az informatikában manapság az eredetét tekintve másik, de hasonló eredményekkel járó jelenség tapasztalható. Ott újra fölfedezik a tezauszokat, és még a jobbik eset, hogy nem adnak neki valamilyen más nevet (pl. fogalomtár). Olyan beszámolók jelennek meg, amelyek szerzői szemmel láthatóan nem tudják, hogy ezeket az információkereső-nyelvi szótárakat már a múlt század 50-es éveitől kezdve használják [12][18].
- 2 Összesen 6 magyar nyelvű tezauszszabvány készült eddig [5][6][7][8][9][11]; értékesebbnél értékesebb információkat tartalmaznak a szerkezetről és a készítésről.
- 3 A magyarázó relációneveket az archívumot látogató, humán műveltségű filmes felhasználókra tekintettel alkalmazzák [1]. A cég könyvtári keresésben járatlan üzletkötői miatt használnak magyarázó relációneveket például az *Arany Oldalak* szakmai telefonkönyv háttértezauszában is [10]. Az itt említett *Filmtezausz* és *Szakmatezausz* megtekinthető az OSZK Könyvtári Intézetének szakkönyvtárában, ahol a legtöbb Magyarországon készült tezauszhoz – 36 db – is létezik példánya. Raktári jelzetük a [16] irodalomjegyzékéből megállapítható.
- 4 A tezauszoknak, tárgyszó-rendszereknek, sőt az osztályozási rendszereknek is létezik már nemzetközi és nemzeti (magyar) adatcsere-formátuma, és a

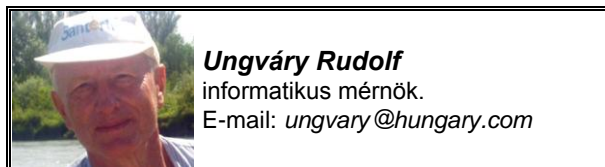
könyvtári rendszereket arra készítik föl, hogy az ilyen formátumban készült (és **csak** az ilyen formátumban készült) információkereső-nyelvi szótárakat tudják importálni [2].

- 5 A HUNMARC 001 hívójelű adatmezője.
- 6 A Relexszel eddig kb. 30 tezausz készült el, és számos könyvtárban, különösen a Köztauruszt használó könyvtárakban a katalogizáláskor profeszszionálisan használják annak érdekében, hogy a legtalálhatóbb kifejezést gyorsan megtalálják, mert könynyebb benne navigálni, mint magában a könyvtári rendszer katalogizáló moduljában [14].
- 7 Még a magyar nyelvű újabb szakirodalomban is rendelkezésre állnak azok a források, amelyek alapján az információkereső nyelvekkel és fogalmi rendszerekkel (többek között tezauszokkal, tárgyszó-rendszerekkel, mi több: ontológiákkal megismerkedni kívánó, eljövendő szakmai élete iránt igényes) – és nem utolsósorban eladható tudásra törekvő – kezdő szakember tájékozódhat alapvető elméleti ismeretekről [13][15 [17][18][19].
- 8 A fogalom nem nyelvfüggő. Előfordul, hogy adott nyelven nincsen rá egyszerű vagy összetett szó (ilyen eset például a német „Backen” és „Braten”, melyet magyarul egyaránt a „sütés” kifejezéssel nevezünk meg), és az is gyakori, hogy a megnevezés hangalakja, vagy asszociációja miatt két nyelv szava szigorúan véve nem pontosan ugyanazt a fogalmat nevezi meg (például a magyar „tisztítás”, „tisztántartás” fogalmába kevésbé játszik bele érdemben a lelki megtisztulás és az erkölcsös fogalma, ezzel szemben az angol „purification” megnevezésbe igen. Ezért a „tisztítás” és a „purification” nem teljesen azonos fogalmakat neveznek meg, de az ilyen árnyalatnyi különbségeknek inkább csak a szépirodalmi szövegekben van jelentősége (noha – különösen a politikai – közbeszédben is okozhatnak bizonyos kifejezések adott esetben problémákat az áthallás miatt).
- 9 Ha szinonimáról van szó (pl. kutya és eb), akkor a két különböző név ugyanazt a fogalmat azonosítja, ha kváziszinonimáról, akkor a két név két különböző fogalmat azonosít, melyet az adott rendszerben öszszevontan kezelnek a deskriptorral (pl. a 'vadászku-tya' nemdeskriptor helyett a 'kutya' deskriptor használandó).
- 10 Az angolban is létezik a magyar F/A generikus relációnak megfelelő BTG/NTG (broader term generic/narrower term generic) reláció (2. táblázat)!
- 11 A megjegyzések jelölését a tezauszszabvány írja elő. A korabeli MARC adatcsere-formátum ismeretének hiányában a nemzetközi és a nemzeti tezauszszabványokban sajnos csak egy megjegyzést szabványosítottak [10][11][17].
- 12 Utoljára 1968-ban SZILÁGYI Tibor foglalta össze a „Tárgyszó és a tárgyszókatalógus” c. művében az összetett tárgyszavakra vonatkozó tudnivalókat.

Irodalom

- [1] Filmteszaurusz. A Magyar Nemzeti Filmarchívum teszaurusza: 3.0 változat. Szerkesztett Ungváry Rudolf. Budapest, MNA, 2006. XVIII, 565 p.
- [2] A MARC 21 Format for authority data. 1999 ed. Update No. 1 (October 2001) through Update No. 8 (October 2007). Washington, Library of Congress, Network Development and MARC Standards Office. <<http://www.loc.gov/marc/authority/ecadhome.html>> A MARC 21 Format for classification data. 2000 ed. Update No. 1 (October 2001) through Update No. 8 (October 2007). Washington, Library of Congress, Network Development and MARC Standards Office. <<http://www.loc.gov/marc/classification/eccdhome.html>> MARC21. Handbuch in deutscher Sprache. Updated November 2008/yb. [Bern], Schweizerische Nationalbibliothek, 2008. <<http://ead.nb.admin.ch/web/marc21/dmarcein2.htm>> Formato IBERMARC para registros de autoridad. [Madrid], Biblioteca Nacional, 1999. ISBN 84-88699-44-1 <http://www.bne.es/docs/Registros_autoridad2001.pdf> HUNMARC, a besorolási rekordok adatcseré formátuma. Összeáll. Sipos Márta, Szabó Julianna, Ungváry Rudolf. [Budapest], Könyvtári Intézet, 2009. Tervezet.
- [3] ISO/DP 5954-1985 Guidelines for the establishment and development of multilingual scientific and technical thesauri for information retrieval.
- [4] ISO/IS 2788-1986 Guidelines for the establishment and development of monolingual scientific and technical thesauri for information retrieval.
- [5] Köztaurusz. 2009. január 1. <<http://www.ki.oszk.hu/107/page.php?57>> és <<http://mek.oszk.hu/adatbazis/thes.htm>>
- [6] MSZ 3418. Magyar nyelvű információkereső teszauruszok szerkezete, részei és formái. Budapest, MSZH, 1998. p. 10.
- [7] NTP 3-75 Többnyelvű információkereső teszaurusz. Tartalom, szerkezet és megjelenési forma.
- [8] NTP 9-75 Az egynyelvű információkereső teszaurusz kidolgozásának szabályai.
- [9] NTP 15-76 A többnyelvű információkereső teszaurusz kidolgozásának szabályai.
- [10] UNGVÁRY Rudolf: Szakma teszaurusz. 5.3 változat. [Közl. a Magyar Telefonkönyvkiadó Társaság]. Budapest, Ariel Studio Bt, 1999. VII, 78. KMK: 4-9870
- [11] SZT SZEVI 174-75 Egynyelvű információkereső teszaurusz. Szerkezet, összetétel és megjelenési forma.
- [12] UNGVÁRY Rudolf: A teszauruszok felfedezése. Eljuttatni a definícióig. = Új Alaplap, 13. köt. 8. sz. 2000. aug. p. 62–63. <<http://www.ontologia.hu/document/paper/>>
- [13] UNGVÁRY Rudolf – ORBÁN Éva: Osztályozás és információkeresés. Kommentált szöveggyűjtemény. Budapest, OSZK, 2001. 2 db.
1. köt., Az osztályozás és elmélete. 543 p. ISBN 963 200 425 6
2. köt., Az információkeresés és elmélete. 535 p. ISBN 963 200 426 4
- [14] UNGVÁRY Rudolf: Tezauruszkezelő programok és a RELEX. = TMT, 48. köt. 1. sz. 2001. p. 3–16. <http://tmt.omikk.bme.hu/show_news.html?id=1620&issue_id=26>
- [15] UNGVÁRY Rudolf: A teszaurusz a felhasználói felületen. Az optimális teszauruszmegjelenítés problémái adatbázis-kezelő rendszerekben. = TMT, 48. köt. 3. sz. 2001. p. 3–12. <http://tmt.omikk.bme.hu/show_news.html?id=1620&issue_id=26>
- [16] UNGVÁRY Rudolf: Az OSZK teszaurusz és a KÖZTAURUSZ. = Könyvtári Figyelő, Új folyam 11. (47.) köt. 1. sz. 2001. p. 11–40. <<http://primula.oszk.hu/kiadvany/kf/2001/1/ungvary.html>> <<http://www.oszk.hu/kiadvany/kf/2001/1/ungvary.html>>
- [17] UNGVÁRY Rudolf: Az információkeresés szavai. = TMT, 50. köt. 12. sz. 2003. p. 1–27. <http://tmt.omikk.bme.hu/show_news.html?id=3451&issue_id=446>.
- [18] UNGVÁRY Rudolf: Tezaurusz és ontológia, avagy a fogalmi ismertetőjegyek generikus öröklődésének formalizálása. = TMT, 51. köt. 5. sz. 2004. p. 175–191. <http://tmt.omikk.bme.hu/show_news.html?id=3615&issue_id=450>
- [19] UNGVÁRY Rudolf – VAJDA Erik: Könyvtári információkeresés. Második, jav. kiad. Budapest, Typotex, 2002. 172 p. ISBN 963 9326 52 8. Kemelen Farkas Digitális Könyvtár <<http://www.hik.hu/tankonyvtar/site/books/b153/index.html>>
- [20] Nemzeti könyvtárak besorolási adatállományai: Library of Congress Authorities: <<http://authorities.loc.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First>> Recherche autorités RAMEAU: <<http://catalogue.bnf.fr/>> Biblioteca Nacional de España. Autoridades. <<http://catalogo.bne.es/uhtbin/authoritybrowse.cgi>> YSA. Yleinen suomalainen asiasanasto: <<http://vesa.lib.helsinki.fi/ysa/index.html>>

Beérkezett: 2008. XI. 14-én.



Ungváry Rudolf
informatikus mérnök.
E-mail: ungvary@hungary.com