

Ungváry Rudolf

Ismeretszervező-könyvtári rendszerek tartalmi feltárásának összehasonlító vizsgálata MARC21 környezetben

Király Péter összevetette a legjelentősebb nemzeti könyvtárak bibliográfiai MARC21 állományát. E tanulmányban csak a források tartalomra vonatkozó adatait és a velük összefüggő adatelemeket elemeztük, összehasonlítva a Deutsche Bibliothek, a Genti Egyetemi Könyvtár, az MTA Könyvtára és a Szegedi Tudományegyetem Könyvtárának adatait. A problémákat nem a MARC21, hanem annak gyakorlati alkalmazása okozza. A MARC21 tartalmi adatai lehetőségei messze nincsenek kihasználva. A könyvtáros felhasználó számára érthető és kezelhető logikai adatcsere-formátumra mindig szükség van és lesz. Az igazi problémát a MARC-ot kezelő katalogizálási szoftverek okozzák.

Tárgyszavak: könyvtári katalógus; katalogizálás; szabályzat; osztályozásmélt; könyvtárügy; szabvány

1. Bevezető. A tanulmány forrása és tárgya

Király Péter letöltötte és formálisan részletesen elemezte 16 jelentős nemzeti és tudományos könyvtár (köztük két magyar nagykönyvtár és a MOKKA) bibliográfiai MARC21 [11] állományát [6]–[10].¹

Ezek a következők:

Baden-Württembergi Közös Katalógus (BZB),
Bajor Közös Katalógus (BAY),
Kézi Nyomtatású Könyvek Adatbázisa (CERL),
Német Nemzeti Könyvtár (DNB),
Európai Tudományos Könyvtárak könyvtörténeti adatbázisa (CERL),
Finn Nemzeti Könyvtár (NFI),
Genti Egyetemi Könyvtár (GENT),
Hannoveri Műszaki Könyvtár (TIB),
Harvardi Egyetemi Könyvtár (HAR),
Kongresszusi Könyvtár (LOC),
Michigani Egyetem Könyvtára (MIC),
MOKKA,
Nemzetközi Zenei Források Repertórium (RIS),
Magyar Tudományos Akadémia könyvtára (MTAK),
San Franciscói Közművelődési Könyvtár (SFP),
Stanfordi Egyetem Könyvtára (STA),
Szegedi Egyetemi Könyvtár (SZTEK),
Torontói Közművelődési Könyvtár (TOR).

Az OSZK katalógusállományának letöltéséhez nem járultak hozzá.

A webes felületre csak hat, a fenti felsorolásban félkövéren írt katalógus és a MOKKA feldolgozása került, mert a göttingeni szerveren nem volt több erre a célra rendelkezésre bocsájtott hely. Még az is megjegyzendő, hogy vannak olyan szervezetek, melyek kifejezetten szeretnék a saját adataik statisztikáit összevetni másokéval, mások pedig ezt nem szeretnék – az OSZK tehát nincs egyedül. Meg kell becsülni azoknak a könyvtáraknak a gyakorlatát, amelyek nyitottak a közzétételre², mert ez szakmailag rendkívül hasznos. Megjegyzendő, hogy a DNB-ben nem a MARC21-et, hanem a PICA formátumot használják és a DNB konvertálta át MARC21 formátumra.

Az elemzés kiterjedt a MARC21 (és adott esetben HUNMARC) összes mezőjére és almezőjére (azaz az összes adatelem-típusra és -értékre). Ezáltal többek között rendelkezésre áll az egyes mezőkben, illetve almezőkben a rögzített adatelem-értékek száma és az, hogy ez hány százaléka a katalógus összes mezőjében szereplő értéknek. Megtalálható az értékek állománya és a rekordok, melyekhez az egyes értékek tartoznak. Az 1. ábrán³ a 650-es mező első néhány almezőjének elemei (almezőjel, név stb.) láthatók (nem az összes).

path	label	terms	records	
			count	%
Numbers and Code				
650 — Subject Added Entry - Topical Term				
\$1			9	0.00%
\$2	Source of heading or term		733,857	57.82%
\$3	Materials specified		1,922	0.15%
\$a	Topical term or geographic name entry element		733,809	57.81%
\$b	Topical term following geographic name entry element		26	0.00%

1. ábra A 650-es mező első néhány almezője

A „path” (=hely) oszlop kék almező-hívójelre kattintva megjeleníthetők azok a bibliográfiai rekordok, melyekben az almező előfordul, a következő „label” oszlopban következnek a mező neve és az almezőinek neve, a „terms” (=értékek, melyek lehetnek számok, jelzetek [„kódok”⁴], természetes nyelvű kifejezések és szövegek) oszlopban a kék számozott felsoroló ikonra kattintva az almező tartalma jeleníthetők meg, és ha ezek egyikére kattintunk, megjelenik a rekord, amelyben az adott érték szerepel. Ennek lekérhető akár a teljes, akár a MARC21 formátuma. A „count” mutatja az adott

almezőben rögzített rekordok számát, az utolsó oszlopban szerepel, hogy ezek száma hány %-a az összes rekord számának. Az almezőnév után, a harmadik oszlopban grafikusan szerepel szemléltetésül a %-os arány. Ha a százalékos arány kerekített értéke 0,01-nél kisebb, egységesen 0,00 szerepel.

A 00X mezőcsoportban az egyes pozíciókban, ahol elvileg csak előírt (kötött szótárból származó) értékek lehetnek, ki vannak mutatva az értékek (2. ábra).

path	message	url	instances	records
control subfield level issues			3,656,937	363,434
control subfield: invalid code (158 variants) [+]			4,807	4,556
008/18-21 (tag008book18)	'r' in 'ar '	i	3,392	3,392
008/18-21 (tag008book18)	'r' in 'arb '	i	292	292
008/18-21 (tag008book18)	'r' in 'r '	i	45	45
008/18-21 (tag008book18)	'r' in 'arf '	i	28	28
008/18-21 (tag008book18)	'r' in 'arbc'	i	25	25

2. ábra A 00X mező egyik hibakimutatásának részlete

A kék [+] parancsjelre kattintva megjelennek azok a pozíciók, amelyekben hiba van, az utána következő „message” (=„üzenet”) oszlopban a konkrét hibák, a kék i parancsjelre kattintva a MARC21-formátum vonatkozó oldalai, majd következnek az előfordulások számai („instances”) és végül a rekordok számai („records”).

- a) Az elemzés alapján következtetések vonhatók le egyrészt arról, hogy milyen főbb feldolgozási jellegzetességei vannak az egyes könyvtárak katalógusainak:
- milyen kimerítően alkalmazzák a MARC21 mezőit és almezőit (használgják-e például a 648 kronologikus tárgyító mezőjét);
 - mennyire kitöltöttek az almezők, (például mennyire fordul elő kitöltött \$0 besorolási rekordazonosító almező vagy a \$1 URI⁵ almező);
 - milyen a koherencia összetartozó almezők között (például milyen az arány a \$a rendszó és a \$2 forrás kitöltöttsége között) stb.;
 - mennyire alkalmazzák a helyi (lokális) célra fenntartott 9XX mezőket és a \$9 almezőket;
 - mennyire fordulnak elő a MARC21-ben nem szereplő, (magyar könyvtárak esetében a HUNMARC-ban sem szereplő) mezők, almezők.
- b) Az elemzés alapján következtetések vonhatók le az alkalmazott könyvtári feldolgozórendszerrel (például mennyire engedi meg a hibás értékek bevitelét), továbbá az alkalmazott kötött

szótárakról (tárgyszórendszer, teaurusz, osztályozási rendszer, taxonómia).

- c) Végül pedig, de nem utolsósorban összehasonlíthatók az egyes könyvtárak katalógusai egymás között a feldolgozás szempontjából (például hol használják, illetve nem használják a 034 mező geokoordinátáit, a 653 szabadon fölvehető tárgyszó mezőt).

Teljes körű következtetéseket valójában csak az adott könyvtár szakemberei képesek levonni, mivel lehetnek olyan esetek, hogy a jellegzetességnek, hibának az okait csak ők tudhatják. A tanulságok azonban a szakma közös ügyei.

A tanulmányban csak a források *tartalomra vonatkozó* adatelemeit (043, 044, 045, 052, 072, 080, 082, 084, 085, 6XX) elemeztük, és velük összefüggésben néhány más adatelemet (008, 034, 041). Csak a Német Nemzeti Könyvtár, a Genti Egyetemi Könyvtár, a Magyar Tudományos Akadémia Könyvtára, a Szegedi Egyetemi Könyvtár és a MOKKA adatait hasonlítottuk össze.

Azok az almezők, melyek csak a HUNMARC-ban léteznek, halványak, azok, melyek se a MARC21-ben, se a HUNMARC-ban nem szerepelnek, halványak és az almezőjel után három kérdőjel van. Azok az értékek, melyek valamilyen okból figyelemre méltók, félkövérek.

A tanulmányban mezőnevek dolgában a HUNMARC 2002. március magyar KSZ/4.1 kiadás terminológiáját követem [5]. Ha a MARC21-ben eltérés van, vagy az almező új, akkor az angol kifejezés fordítását adom meg.⁶

2. A vezérlő mezők és előírt formátumú almezők hibáinak összehasonlítása

Király Péter elemzésében az „Issues” („Hibatáblázatok”) alatt kimutatja a 00X-es mező hibáit is. Ebben a mezőben csak néhány karakteres értékek vannak, melyek kötött szótárból származnak, és a mező pozícióhoz kötve kell őket rögzíteni (00X). Ugyancsak megjelennek az előírt számformátumú

almezők (pl. ISBN) és az előírt formátumú jelzet almezők (pl. \$6 és hasonló) hibái is.

A 008-as, 40 karakter hosszú, pozícionált adatelem-típusokra felosztott mező hibái különösen feltűnőek. E mező értékei *szorosan összefüggnek* a tartalom szerinti kereséssel.⁷ Az állandó hosszúságú, pozícionált értékek ugyanis a dokumentumok típusát, műfaját megadó néhány karakteres jelzetek, melyek a 6XX mezőcsoportok tárgyszavaival együtt (de elvileg bármilyen más mező értékeivel együtt) finomíthatják a keresést. Más szóval keresés során a nagyon is fontos szűkítési szempont a dokumentum típusa. E jelzetek jelentőségéhez képest a mező nem igazán kapja meg az értékeit megérdemlő jelentőséget, még ha az egyes pozíciókon megadható formai jelzetek nagyon korai időszakban (a 70-es években) készültek, és osztályozásméletileg nem tekinthetők igazán koherens rendszereknek (1. táblázat).

A pozícióhoz kötött formai jelzeteken *helyileg* olykor ugyan módosítottak⁸, de magában a MARC21-ben szereplő tipológia a régi maradt, a lényegi, és osztályozásméletileg korrektebb átszerkesztésükre nem került sor. A legnagyobb probléma azonban velük az, hogy a könyvtári kezelőrendszerek nem biztosítják igazán a végfelhasználó számára egyszerűen kombinálható keresést.

A 00X mezőkben pontosan meg van határozva a számérték (pl. dátum), illetve a jelzet („kód”) formája, karakterei. A pozícionált jelzetek („kódok”) ugyanis kötött szótárból származnak, melyek a MARC21 részei. Ehhez képest a hibák száma és fajtája szembeszökően nagy. És ez nem korlátozódik egyes könyvtárakra, hanem általános. Valószínű, hogy itt nem elsősorban a feldolgozás gondosságáról, minőségéről van szó, hanem a katalóguskezelő szoftver hibaellenőrző képességei vannak aluldefiniálva, még hozzá eléggé. Az ilyen típusú, nem természetes nyelvű értékek még oly figyelmes intellektuális rögzítése sem képes kizárni a hibákat.

1. táblázat.

Pozíció-, mező- és almezőhibák (a 2. ábrán példaként a 008-as mező 18–21. pozícióin a hibás jelzetek megjelenítése látható)

MARC21 mező/adatelem	DNB ért.	DBN rek.	GENT. ért.	GENT rek.	MTAK ért.	MTAK rek.	SZTEK ért.	SZTEK rek.	MOKKA ért.	MOKKA rek.
008 meghatározott jellemzők és információs adatok										
egyéb vezérlő mezők és almezők										
nem ismételtető mező ism. ⁹	740	740	65	67	8	8	240	240	4705	4698
nem előírt hívójelű mező ¹⁰	–	–	–	–	308 647	268 696	2 884	2 626	1 115 714	479 623
\$ind hibás indikátor	53 161	8 066	3 031	5 066	192 690	178 461	5 015	4 271	3 451 444	2 298 715
\$X hibás almező ¹¹	2 368 900	2 483 725	430 833	529 103	87	57	31 197	26 900	34 759 334	8 864 772
ebből 65X hívójelű almező	728	768	16	16	21	21	75	71	682	624
nem ism. almező ismétlődése	–	–	–	–	6 452	3 867	1 900	1 857	881 787	776 565

Az 5. sorban a „nem ismételtető mezők ismétlődése” különösen szemet szúr: formálisan nagyon egyszerű hibatípusról van szó.

A 6. sorban a nem előírt hívójelű mezők és almezők egész kis részét lehet, hogy helyi használatra vették föl, többségükben az értékek annyira következtelen (zajos) képet mutatnak, hogy ezek az almezők csak meglehetősen kis valószínűséggel lehetnek tervezettek. A 65X mezők hibái elsősorban beviteli hibák lehetnek. A két magyar tudományos könyvtár esetében pedig részben HUNMARC hívójelekről és almezőjelekről lehet szó, melyek almezőneve nem került be a Király Péter által készült szoftverbe. Az is lehet, hogy az ilyen hibák más rekordok betöltése során keletkeztek.

Ezeknek az adatelemeknek az értékei, továbbá a mezőhívójelek és almezőjelek, valamint az előírt egy karakteres indikátorok olyan szerkezetűek, hogy formálisan viszonylag egyszerűen ellenőrizhetők; ezért ellenőrzésük a kezelőrendszerekbe nehézségek nélkül beépíthetők lettek volna. Nem megengedett pozícióértékek, nem ismételtető mező ismétlődésének, nem előírt hívójelű mezők, almezők, indikátorok bevitelének stb. megakadályozása vagy utólagos kiszűrése aligha ördögös feladat. Mégis, még olyan nagy reputációval rendelkező könyvtárakban is olykor milliós az ilyen hibák nagyságrendje (mint például a DNB közel két és fél millió nem előírt hívójelű almezője).

A MOKKA közös katalógusba bizonyára szűrést követően történt a letöltés. Az OSZK hibáinak ki-

szűrését Szabó Julianna végezte el. Lásd: „MOKKA inicializálás. Az OSZK rekordok szintaktikai ellenőrzése. Egykötetes és többkötetes közös adatokat leíró rekordok. 2003. 02. 07. – 92 p. [kézirat]” A MOKKA honlapján (<http://www.mokka.hu/web/guest/katalogizalasi-szabalyok>) ezeknek nincs nyoma, csak a katalógizálási szabályzat és a Szabó Julianna által készített segédlet található.

Feltűnő, hogy a többi könyvtárból mennyi hiba került át. Akik a kérdést jól ismerik, tudják, hogy a szűrőprogramok, valamint a lokális mezők adminisztrációja, leírásuk dokumentációja általában nem valami tökéletes, és ami van, nehezen hozzáférhető, és ez a MOKKA esetében is így lehet. Egyedül az OSZK által átadott állomány szűrés specifikációja ismert.¹⁵ A MOKKA honlapján szerepelnek a specifikáció dokumentumai, de ezek között nincs olyan, mely az egyes könyvtárak állományának konvertálási specifikációját tárgyalja.¹⁶

Célszerű volna, ha a könyvtári feldolgozórendszerek a lokális mezők és almezők definiálásakor lehetőséget biztosítanának arra, hogy ne csak megnevezésük legyen ezeknek a mezőknek és almezőknek. Ezek ugyanis tömörségük miatt nem eléggé informatívak. Hosszabb magyarázat, meghatározás is kapcsolható legyen hozzájuk, melyeket a rendszer nyilvántart, és opcionálisan felkínálja őket.

Mutatóban néhány hiba az SZTE állományából, melyekhez hasonlók más könyvtárakban is előfordulnak (2. táblázat).

2. táblázat
Hibapéldák

<i>hibatípus:</i>	<i>hiba:</i>	<i>rekord db:</i>	<i>előfordulás db:</i>
Hibás 008-as mező (720 változat)		3 656 937	363 434
Hibás jelzet (158 változat):		4 807	4 556
Pl. 008/18-21	'r' in 'ar'	3 392	3 392
<i>Azaz a könyvek 18–21. pozícióin az egy karakteres 'a' (illusztráció) érték helyett hibásan 'ar' szerepel.</i>			
Hibás érték (562 változat)		3 652 130	359 259
008/00-05	'2002'.	13	13
<i>Azaz a könyvek 00–05. pozícióin a helyes dátumforma ÉÉHHNN.</i>			
008/29	'[köz]'	225 388	225 388
<i>Azaz a pozíción '[köz]' nem fordulhat elő.</i>			
008/06	'x'	1	1
<i>Azaz pozíción 'x' nem fordulhat elő.</i>			
Hibás indikátor		333	332
65\$ind1	1	33	33
<i>Az első indikátor helyén 1 szerepel, holott térköznek kell lennie.</i>			
Hibás almező		31 197	26 900
080\$6	hibás karakterek az almezőben: '.', '')	8	8
<i>Azaz az ETO-jelzet \$6 írásrendszer-jelzetét tartalmazó almezőjében csak a MARC21 „A függelékében” szereplő \$6 almezőhöz megadott kötött szótár hat megengedett írásrendszert meghatározó jelzet közül lehet választani (pl. ' (N\$a ' = cirill írás), ennek ellenére pont, vessző és záró kerek zárójel fordul elő a rögzített jelzetben, mely karakterek a kötött szótár egyetlen jelzetében sem szerepelnek.</i>			
Almező: hibás ISBN (10348 változat)		13 115	13 012
020\$a	A '0238-3349(fűzött)' nem ISBN-formátum	8	8

Összegezve: a túlnyomórészt előírt formátumú számértékeket vagy egyszerű, kötött szótárból származó jelzeteket tartalmazó mezők hibái megelőzhetők lennének, ha maguk a könyvtári rendszerek lennének jobban felkészítve ezek ellenőrzésére, hibaüzeneteire. Ehhez természetesen szükséges, hogy a feldolgozó könyvtárosok is közvetíteni tudják a pontos hibaellenőrzési igényeiket a kezelőrendszerek beszerzésekor és telepítésekor. Jó volna, ha erre az oktatásban is felkészülnének.

3. A 0XX számozások és jelzetek mezőcsoportja

A mezőcsoport különféle rekordazonosítóinak és a 040 katalógizálás forrásának a kitöltöttsége képet

ad az egyes könyvtárakban folyó feldolgozás teljességéről. Ezen túlmenően tartalmazza a nyelv és az ország azonosítóit („kódjait”), melyeknek szerepe van a tartalom szerinti keresésben.¹⁷ Ugyancsak fontos szerepe van földrajzi források esetében a 034 mező geokoordinátáinak (melyek a jövőben a földrajzi nevek alapján a weben a térképi megjelenítés lehetőségét biztosítják).¹⁸ Végül a mezőcsoportban hét tartalomfeltáró mező is szerepel, melyek többsége speciális osztályozási rendszerek jelzeteit tartalmazza, de van közöttük tárgyszavas rendszerek mezője is. Mindez indokolta, hogy az összehasonlító elemzésembe a mezőcsoport mezőinek többsége bekerüljön.

3.1 A bibliográfiai rekordazonosítók (3. táblázat)

3. táblázat

A OXX mezőcsoport bibliográfiai rekordazonosítói

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
010\$a LC azonosító	5894	0,03	256 940	14,56	1	0,00	161	0,01	56 366	0,63
015 nemzeti bibliográfiai azon-										
\$2 forrás	11 345 731	59,63	26 219	1,18	–	–	4	0,00	593	0,01
\$a azonosító	11 345 731	59,63	71 517	1,49	–	–	7	0,00	6 284	0,07
\$z törölt azonosító	1 995 499	10,49	637	0,04	–	–	–	–	9	0,00
016 bibl. intézményazon.										
\$2 forrás	19 mill,	100	31 802	1,80	–	–	5	0,00	840	0,01
\$a azonosító	19 mill,	100	32 155	1,82	–	–	6	0,00	844	0,01
019\$a [helyi adatelemtípus]	–	–	2	0,00	–	–	–	–	–	–

Érthető, hogy egy nemzeti könyvtárba (DNB) alig kerül más nagykönyvtárból származó rekord, mely a 010\$a LC-azonosítót tartalmazza. Ezért a DNB 0,03%-nyi LC-azonosítót tartalmazó rekordja igazi, de nagyon kevés kivétel. A GENT-ben viszont láthatóan több ilyen rekord is van (14,56%). Az tehát magyarázatra szorulhat, hogy vannak tudományos könyvtárak, melyekbe elenyésző mértékben vannak jelen az LC-ből származó, vagy annak rekordjaival azonos rekordok, holott feltételezhető, hogy ezekbe akár nagyobb számban is átvesznek ilyen rekordokat.

A 016 nemzeti bibliográfiai intézmény azonosítójának a kitöltöttsége a német nemzeti könyvtár esetében érthető módon teljes. Ez abból is látható, hogy értékeinek száma (19 millió) megegyezik a 4. táblázatban látható 040 *katalogizálás forrása* mező értékeinek a számával – ami egyébként összehasonlításban is más könyvtárakkal nagyon gondos

feldolgozásra vall. Szorgos feldolgozói tevékenységre utal a törölt azonosítók nagy száma is.

Más könyvtárakban a saját 016-os bibliográfiai azonosítói alig jelennek meg, és helyi 019-es azonosítók szinte nincsenek. Hasonló a helyzet a 015-ös nemzeti bibliográfiai azonosítóval. A DNB esetében csak a katalógusrekordok 59,63%-a kerül be a nemzeti bibliográfiába. A – Király Péter elemzésében nem szereplő – OSZ esetében kb. 65% a könyveknél a bekerülési arány, de ez évenként ingadozhat. A többi könyvtár gyakorlatilag nem

tüntetni föl ezt az azonosítót sem. A mégis bekerülő féltucatnyi rekord oka talán a pusztá feldolgozói véletlen, vagy betöltött rekordok következménye. Megjegyzésre érdemes, hogy Király Péter elemzett állománya szerint legalábbis a MOKKA nem vette át az OSZK-tól egyik bibliográfiai azonosítót se. Minderre hozzáértők számára nemcsak érdekes lenne, de szakmailag meglehetősen hasznosak lehetnének az adott könyvtárak szakembereinek magyarázatai, értelmezései.

3.2 A kiadványazonosítók

A két kiadványazonosító a se a MARC21-ben, se a HUNMARC-ban nem szereplő, ismeretlen almező-jelek és ezek almezőtartalmait illetően afféle állapot-orvosi ló szerepét játssza. A hibát tartalmazó rekordok abszolút száma ugyan elhanyagolható nagyságú, tehát nem vonható le belőlük az érintett katalógusok minőségéről általános következtetés, annál inkább ad lehangoló bizonyítványt a katalogizáló szoftver képességeiről, ami a hibalehetőségek ellenőrzési képességeit illeti (4. táblázat).

Ilyen zaj a külföldi könyvtárakban meglehetősen ritka. Valószínűsíthető, hogy elsősorban a szoftver (a könyvtári feldolgozórendszer) minősége ennek oka. A jobb minőségű szoftver hiánya meg sajnbajna elsősorban anyagi. Ami miatt nemcsak a szoftverek gyengébb definiáltságúak, és nincs pénz utólagos fejlesztésekre, de a felkészült szakembereket se tudja a távozástól visszatartani a pénzügyi szűkösség – ami viszont már régi (kultur)politikai probléma. A „nemzet csinósítása” alulfinanszírozott.

4. táblázat
Az ISBN és az ISSN

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
020 ISBN										
\$- ???	-	-	-	-	-	-	-	-	2	0,00
\$0 ???	-	-	-	-	-	-	25	0,00	45	0,00
\$1 ???	-	-	-	-	-	-	9	0,00	139	0,00
\$2 ???	-	-	-	-	2	0,00	18	0,00	79	0,00
\$3 ???	-	-	-	-	1	0,00	21	0,00	91	0,00
\$4 ???	-	-	-	-	2	-	3	0,00	29	0,00
\$5 ???	-	-	-	-	-	-	3	0,00	30	0,00
\$6 kapcsolódás	-	-	68	0,00	2	0,00	12	0,00	531	0,01
\$7 ???	-	-	-	-	1	0,00	1	0,00	16	0,00
\$8 mezőkapcs, és soroz,	-	-	3	0,00	-	-	6	0,00	35	0,00
\$9 [helyi almező]	5 759 980	30,27	-	-	-	-	91	0,01	22	0,00
\$C ???	-	-	-	-	-	-	-	-	3	0,00
\$a ISBN azonosító	5 721 373	30,07	621 661	35,24	217607	20,68	645 528	50,86	27 385 51	30,81
\$b ???	-	-	-	-	2	0,00	23	0,00	457	0,00
\$c hozzáférés feltétele ¹²	7 044 386	37,02	9 950	0,56	45	0,00	25 153	1,98	30 041	0,34
\$d árváltozatok	-	-	-	-	-	-	28	0,00	971	0,01
\$e ???	-	-	-	-	-	-	1	0,00	12	0,00
\$f ???	-	-	-	-	-	-	-	-	2	0,00
\$g ???	-	-	-	-	-	-	55	0,00	57	0,00
\$h kiegészítő információ	-	-	-	-	-	-	2	0,00	696	0,01
\$i megjegyzés	-	-	-	-	-	-	25	0,00	3 880	0,04
\$j kötés	-	-	-	-	-	-	283	0,00	702	0,01
\$k ???	-	-	-	-	-	-	-	-	1	0,00
\$l ???	-	-	-	-	-	-	2	0,00	2	0,00
\$n ???	-	-	-	-	-	-	1	0,00	-	-
\$q minősítő információ	-	-	87 028	4,93	8 095	0,77	6	0,00	4 524	0,05
\$r ???	-	-	-	-	-	-	-	-	2	0,00
\$s ???	-	-	-	-	-	-	3	0,00	4	0,00
\$t ???	-	-	-	-	-	-	2	0,00	1	0,00
\$z törölt ISBN	85 751	0,45	8 869	0,50	1 187	0,11	4 283	0,34	47 795	0,54
022 ISSN										
\$- ???	-	-	-	-	1	0,00	-	-	-	-
\$0 ???	-	-	-	-	3	0,00	2	0,00	21	0,00
\$1 ???	-	-	-	-	2	0,00	1	0,00	7	0,00
\$2 forrás	-	-	42	0,00	2	0,00	-	0,00	3	0,00
\$3 ???	-	-	-	-	1	0,00	-	-	-	-
\$4 ???	-	-	-	-	1	0,00	-	-	1	0,00
\$5 ???	-	-	-	-	-	-	1	0,00	-	-
\$6 kapcsolódás	-	-	-	-	1	0,00	-	-	1	0,00
\$7 ???	-	-	-	-	1	0,00	-	-	-	-
\$8 mezőkapcs, és sorsz,	-	-	-	-	-1	0,00	-	-	-	-

Ungváry R.: Ismeretszervező-könyvtári tartalmi feltárásának összehasonlító ...

\$S ???	-	-	-	-	1	0,00	-	-	-	-
\$a ISSN azonosító	1,651 mill,	8,68	427	0,02	19 357	1,84	69743	5,49	85055	0,96
\$b ???	-	-	-	-	-	-	-	-	1	0,00
\$c ár	1343	0,01					28	0,00	872	0,01
\$f ???	-	-	-	-	-	-	3	0,00	-	-
\$g ???	-	-	-	-	-	-	1	0,00	5	0,00
\$i megjegyzés	-	-	-	-	30	0,00	-	-	67	0,00
\$j kötés	-	-	-	-	-	-	-	-	83	0,00
\$I ISSN-L	40 582	021	50	0,00	6	0,00	4	0,00	2	0,00
\$m törölt ISSN-L	-	-	-	-	-	-	-	-	-	-
\$s ???	-	-	-	-	-	-	-	-	7	0,00
\$t ???	-	-	-	-	-	-	-	-	1	0,00
\$w ???	-	-	-	-	-	-	-	-	59	0,00
\$y hibás ISSN	-	-	-	-	10	0,00	119	0,02	246	0,00
\$z törölt ISSN	42	0,00	15	0,00	47	0,00	212	0,02	149	0,00

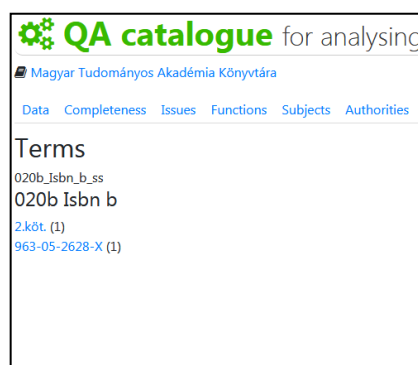
A két magyar könyvtárban és a MOKKA-ban megjelenő, halványan szedett ismeretlen almezők vagy értelmezhetetlen értékeket, vagy sajátos ISBN-illetve ISSN-értékeket, vagy az árat tartalmazzák, miközben a HUNMARC szerint mind az ISBN-ben, mind az ISSN-ben létezik az önálló \$c ár almező is. A MARC21 szerint viszont ISBN esetében a \$c a hozzáférés feltétele almező, az ISSN esetében pedig a \$c a MARC21-ben nem létezik. A két érintett magyar tudományos könyvtár katalogizálási gyakorlatához külső megfigyelő nem rendelkezik elegendő információval, és a sok ismeretlen almezőnek bizonyára lehet magyarázata. A példaként választott külföldi könyvtárak esetében – legalábbis kiadványazonosítók esetében –, ilyen zajnak tűnő jelenséget nem találni.

Önmagában is különös, hogy egyáltalán lehetséges a formátumokban nem létező almezőket rögzíteni és ezeknek értéket adni anélkül, hogy a rendszer legalább üzenetet ne küldene erről a tényről. A könyvtári feldolgozórendszereknek biztosítani kellene az új mezők és almezők későbbi definiálását a hozzáértő és kellő jogosultsággal rendelkező felhasználók által, miáltal ezek is ellenőrizhetővé válhatnak. Egy biztosnak látszik: a könyvtári feldolgozórendszer megfelelő hibaelőzési képességei esetén ilyen fajta hibákból akkor is lényegesen kevesebb keletkezhetnék, ha anyagilag a nagyon erős revízióra nincsenek is meg az adottságok.

A 3. és 4. ábrán egy-egy nem létező, és inkoherens tartalmú almező példái láthatók.



3. ábra A \$1 almező a SZTE állományában



4. ábra A \$b almező az MTAK

3.3 Néhány további speciális azonosító és a katalogizálás forrása

A speciális azonosítókat tartalmazó almezők közöttve ugyancsak a tartalom szerinti válogatást tennék lehetővé. Ez azonban sajnálatosan messze nincsen igazán felhasználóbarát módon biztosítva a katalogizáló rendszerekben. Nem igazán lehet olyan keresőkérdést egyszerűen megadni a rendszernek, hogy adott tárgyszó vagy tárgyszavak

esetében van-e vagy nincs a találatoknak pl. zenemű kiadói azonosítója? A jövő könyvtári feldolgozórendszereknek (a könyvtári szoftvereknek) a tervezésébe nemcsak nagyobb figyelemmel kell bevonni a konkrét tapasztalatokkal rendelkező feldolgozó könyvtárosokat, hanem őket is fel kell készíteni arra, hogy felismerjék, milyen ellenőrzésekre van szükség mindennapi munkájukban (5. táblázat).

5. táblázat

A speciális azonosítók és a katalogizálás forrása

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
024\$a egyéb szabv. azon.	8,479 mill,	44,56	10 3619	5,87	1 644	0,16	867	0,07	97 762	1,10
028\$a zenemű kiadói azon.	1 408	7,40	7 391	0,42	–	–	–	–	1 292	0,01
030\$a CODEN	6,386	0,03	37	0,00	–	–	12 108	0,95	50	0,00
032\$a postai regisztr. azon.	813	0,00	10	0,00	–	–	1	0,00	1	0,00
034\$d geokoordináták.	55 880	0,29	46	0,00	–	–	–	–	12	0,00
035\$a rendszerazonosító	19 mill,	100	283 252	16,06	774 923	73,64	864 342	68,10	7 062 486	79,45
040 katalogizálás forrása										
\$a eredeti intézmény	19 mill,	100	1 764 209	100	1 052 265	100	1 268 276	100	7 062 997	79,45
\$b katalogizálás nyelve	19 mill,	100	–	–	1 052 286	99,99	972 901	76,65	541	0,01
\$c átíró intézmény	19 mill,	100	–	–	1 052 202	99,99	4,470	0,35	176	0,00
\$d módosító intézm.	19 mill	100	–	–	119 379	11,34	237 527	18,71	277 885	2,56
\$e leírási megállapodások	606 123	3,19	–	–	22	0,00	10	0,00	4	0,00

Figyelemre méltó, hogy egyedül a DNB-ben szerepelnek – a teljes állományhoz képest elhanyagolható mennyiségben – geokoordináta-értékek (034\$d), holott ezek idővel, amikor a katalógustételekben szereplő helynek magából a katalógusból kilépve majd a weben térképi helyként megjeleníthetők lesznek, alapvetőek. A hiányuk ez esetben egyáltalán nem a katalogizáló szoftverből következik.

A SZTE 035\$a rendszerazonosítói OCLC-azonosítók.

A 040-es mezőt minden rekordnak tartalmaznia kell. Ugyancsak elvileg a \$d módosító intézmény és a \$e leírási megállapodások almező kivételével a 040-es mező többi \$a, \$b, \$c almezőjének ugyanannyi értéket kellene tartalmaznia. Hogy ez megvalósítható, azt a DNB-ben létező értékek számszerű összhangja igazolja. A \$c átíró intézmény almező értelmezhető úgy, hogy MARC-formátumban ezt nem kell kitölteni. Nyilván ez magyarázza, hogy van könyvtár, ahol ezt érdemben nem töltik ki (pl. a SZTE), ami mégis keletke-

zik, az betöltés következménye. Az MTAK esetében ez majdnem megvan. A GENT esetében, talán hogy elkerüljék az összhangra törekvés nehézségeit, eleve csak a 040\$a eredeti intézmény adatát adják meg, abból kiindulva, hogy ha valakit érdekel, nézze meg a többi almezőadat az eredeti intézmény katalógusában?

A viszonylag gyengén kitöltött 040\$e leírási megállapodások tartalma szorosan összefügg a 005 rekordfej 18. a bibliográfiai leírás szabálya pozíciójának egy karakteres értékével. Mivel a \$e almező formalizálatlan szöveges almező, a rekordfej értéke a perdöntő jelentőségű abban, hogy a konvertáló program készítői képben legyenek arról, hogy használnak-e központosítási jeleket valamelyik – például a 300 főcím adatmezőben?

Talányos, hogy más könyvtárakban ez az összhang miért hiányzik. Lehet, hogy például a \$b almező kitöltése nem kötelező? Vagy a kezelőrendszer nem tudja ellenőrizni a 040\$b nyelv almező értékeinek egyezését a 008 mező 35–37.

nyelv pozíciók értékeivel? (Persze az is lehet, hogy letöltések révén kerültek a rendszerbe.)

3.4 A nyelv

Ebben a mezőben a két kiadványazonosítóhoz hasonlóan a magyar könyvtárakban nagyon sok az ismeretlen, se a MARC21-ben, se a HUNMARC-ban nem szereplő almező, és értékeik többnyire a nyelvek jelzetei („kódjai”). Lehet, hogy a 041 nyelv

mező esetében is helyileg használt almezőkről van szó (és remélhetőleg dokumentálva vannak, ha külső szemlélő publikusság hiányában nem is tud egykönnyen tudomást szerezni róluk), lehet, hogy tévesen lettek rögzítve. (Vagy itt is lehetséges, hogy letöltött rekordok okozták a „fertőzést”? De ebben az esetben a fogadó katalogizáló rendszer miért nem volt képes ezeket kiszűrni?) (6. táblázat).

6. táblázat

A nyelvi azonosítók

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
041 nyelv										
\$0 ???	-	-	-	-	1	0,00	-	-	1	0,00
\$2 kód forrása	-	-	-	-	-	-	1	0,00	2	-
\$1 ???	-	-	-	-	-	-	-	-	1	0,00
\$3 ???	-	-	-	-	1	0,00	-	-	-	-
\$4 ???	-	-	-	-	-	-	-	-	3	0,00
\$6 kapcsolódás	-	-	-	-	1	0,00	-	-	67 268	0,76
\$7 ???	-	-	-	-	1	0,00	-	-	-	-
\$8 mezőkapcsoló és sorsz.	42 000	0,22			-	-	-	-	-	-
\$a nyelv	15,7 mill.	82,58	101 564	5,76	993 704	94,43	148 840	11,73	4 126 631	46,42
\$b összefoglalás nyelve	-	-	4791	0,27	14 261	1,36	15 930	1,26	67 268	0,76
\$c ???	-	-	-	-	-	-	6	0,00	18	0,00
\$d ének vagy szöveg nyelve	-	-	19	0,27	2	0,00	237	0,02	382	0,00
\$e librettó nyelve	-	-	15	0,00	25	0,00	80	0,01	258	0,00
\$f tartalomjegyzék nyelve	-	-	100	0,01	1 095	0,10	1 680	0,13	10 677	0,12
\$g egyéb járulékok nyelve	-	-	119	0,01	2 439	0,23	3 212	0,25	14 527	0,16
\$h eredeti nyelve	569 679	2,99	12 014	0,68	55 266	5,25	70 562	5,56	367 972	4,14
\$i ???	-	-	-	-	2	0,00	6	0,00	7	0,00
\$j filmfeliratok nyelve	-	-	25	0,00	18	0,00	2	0,00	8	0,00
\$k közvet, fordítás nyelve???	-	-	9	0,00	41	0,00	1	0,00	23	0,00
\$l ???	-	-	-	-	9	0,00	29	0,00	60	0,00
\$m egyéb járulékok nyelve	-	-	9	0,00	11	0,00	1	0,00	7	0,00
\$n eredeti librettók nyelve	-	-	2	0,00	1	0,00	12	0,00	20	0,00
\$p ???	-	-	-	-	1	0,00	4	0,00	4	0,00
\$q ???	-	-	-	-	-	-	-	-	-	-
\$r ???	-	-	-	-	1	0,00	8	0,00	37	0,00
\$s ???	-	-	-	-	2	0,00	3	0,00	15	0,00
\$t ???	-	-	-	-	2	0,00	2	0,00	3	0,00
\$u ???	-	-	-	-	3	0,00	1	0,00	2	0,00
\$v ???	-	-	-	-	-	-	-	-	1	0,00
\$x ???	-	-	-	-	-	-	16	0,00	11	0,00
\$y ???	-	-	-	-	-	-	-	-	1	0,00

Egyetlen katalógusban sincs minden rekord esetében megadva a nyelv, de az feltűnő, hogy a GENT és a SZTE katalógusában nagyon kevés rekord kapott ilyen kódot. A legnagyobb a nyelvi jelzetek („kód”) értékeinek az aránya a DNB-ben, de a DNB nemzeti könyvtár, mely érdemben csak az OSZK gyakorlatával lenne párhuzamba állítható. Tapasztalatom szerint az OSZK-ban ez az arány a DNB-nek megfelelő színvonalú. Már csak ezért is sajnálatos az OSZK állományának a hiánya az elemzésben – nem lett volna ok a szégyenkezésre.¹³ A SZTE gyakorlatában eleve csak a lefordított és többnyelvű dokumentumok esetén adnak értéket a 040\$a almezőnek, lehet, hogy a GENT gyakorlata is ez. Alapvetően a helyi katalogizálási szabályzatok különböző értelmezései magyarázhatják, hogy hol milyen további nyelvi almezőket használnak. A hibás almezők automatizáltan könnyen törölhetők. Valójában a szoftvernek kellene már a bekerülés-kor jeleznie a hibát.

A 040\$a *nyelvi jelzet* mező első három karaktere a MARC21 szerint rendszerint azonos a 008 mezőben a 35–37. *nyelvi jelzet* pozíciókon szereplő

értékkel. Kérdés, mennyire valósul ez meg a gyakorlatban, és ha nem, miért.

3.5 A 0XX mezőcsoport tartalmi mezői a 08X mezők kivételével

A 043–052 mezők értékei kötött szótárakból származó osztályozási jelzetek. A 072 mező értékei ugyancsak kötött szótárból származó tárgyszavak. Ezekre is vonatkozik, hogy ugyancsak a tartalom szerinti keresést és válogatást tennék lehetővé. Például olyan kérdés felhasználóbarát, egyszerű formában föltevését, melyben e mezők bármelyikének az értéke kombinálható volna a 6XX bármely tárgyszavával. Például a 651 mező tárgyszava és mondjuk a 044 jelzete („kódja”). Mondjuk '651\$aPilisborosjenő ÉS 044\$bDE-BW' (ahol 'DE-BW' Baden-Württemberg ISO 3166–2 De szabvány szerinti jelzete). Ez ma még talán nem tűnik igazán reális igénynek. Amikor azonban idővel a mainak sokszorosa lesz az információ a weben, így az online katalógusokban is, a nagyobb zaj kiszűrése a mainál fontosabb lesz (7. táblázat).

7. táblázat

A 0XX mezőcsoport tartalmi azonosítói a 08X mezők kivételével

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE		MOKKA	%
043 földrajzi területazonosító										
\$2 helyi területazon, forrása	–	–	2	0,00	–	–	–	–	–	–
\$a földrajzi területazonosító	–	–	91 420	5,18	–	–	14	0,00	8 126	0,09
\$b helyi területazonosító	–	–	28	0,00	–	–	–	–	2	0,00
\$c ISO kód,	536 976	2,82	84	0,00	–	–	14	0,00	–	–
\$d ???	–	–	–	–	–	–	–	–	3	0,00
\$x ???	–	–	e–	–	–	–	3	0,00	–	–
044 kiadó/előállító ország kódja										
\$9 ??	–	–	–	–	3	0,00	18	0,00	7	0,00
\$a MARC országkód	–	–	–	–	3	0,00	18	0,00	733 150	8,25
\$b részterület helyi azonosítója	–	–	556	0,03	–	–	1	0,00	1	0,00
\$c ISO országkód	18,8 mill,¹⁴	98,76	1328	0,08	–	–	–	–	30	0,00
\$d ???	–	–	–	–	–	–	–	–	1	0,00
\$f ???	–	–	–	–	–1	–	1	0,00	–	–
\$h ???	–	–	–	–	–	–	1	0,00	2	0,00
052 földrajzi osztályozás jelzet, terület										
\$2 forrás	–	–	13	0,01	–	–	–	–	–	–
\$a földrajzi terület jelzete	–	–	185	0,01	–	–	–	–	10	0,00
\$b földrajzi részterület jelzete	–	–	42	0,00	17 625	1,67	–	–	9	0,00

Ungváry R.: Ismeretszervező-könyvtári tartalmi feltárásának összehasonlító ...

072 átfogó szakterület (kategória) jelzete										
\$2 forrás	-	-	5 581	0,32	17 423	0,66	-	-	27	0,00
\$a átfogó szakterület	-	-	5 949	0,34	17 625	1,67	3	0,00	3 423	0,04
\$x átfogó részterület	-	-	453	0,03	-	-	-	-	-	-
090?										
\$6 ???	-	-	-	-	-	-	1	0,00	-	-
\$7 ???	-	-	-	-	-	-	1	0,00	-	-
\$a ???	-	-	-	-	-	-	11 370	16	-	-
\$b ???	-	-	-	-	-	-	16	0,00	-	-
\$i ???	-	-	-	-	-	-	4	0,00	-	-
\$a ???	-	-	-	-	-	-	1	0,00	-	-
\$b ???	-	-	-	-	-	-	1	0,00	-	-

Ezeket a tartalomra vonatkozó jelzeteket, illetve tárgyszavakat tartalmazó mezőket bizonyára csak szakkönyvtárakban használják inkább. A SZTE 090-es mezője helyi használatú jelzeteket tartalmaz; a két \$a és két \$b almező szokatlan, de könnyen javítható hiba.

Figyelemre méltó, hogy a MOKKA-ban milyen sok az – egyébként kötelező – 044\$a *MARC országjelzet* („kód”), holott ennek magyarországi használata nem jellemző. A \$c *ISO országjelzet* („kód”) mező viszont szinte üres, holott Magyarországon ezt kifejezetten használják. Valószínűsíthető magyarázata, hogy 2001-ben új almezőként bevezették a MARC21-be a \$c almezőt az ISO 3166-1 és ISO 3166-2 országjelzetek számára, és ettől kezdve a \$a almezőt a MARC országjelzetek számára tartják fenn. Mivel a MARC21 vonatkozásában – legalábbis ebben az esetben – feltehetően nem volt utánkövetés, az ISO-jelzetek ott maradtak a \$a almezőben, és így kerültek át konvertáláskor a MOKKA-ba.

A DNB-ben bizonyára volt utánkövetés, mert a 044\$c *országjelzet* („kód”) mező szinte az állomány 100%-át tartalmazza. (Az OSZK adatai sajnos hiányzanak, de tudomásom szerint itt megtörtént az utánkövetés.) Kérdés az is, hogy összhang van-e minden könyvtárban a 044\$c *országjelzet* almezők értékei és a 008 mező 15–17. pozícióinak *országjelzet* értékei között. Más szóval olyanok-e a katalógizáló rendszerek, hogy ezt ellenőrzik?

3.5 A 08X tartalmi mezők

A 080, 082, 083, 084 és 085 mezők értékei klaszszikus osztályozási rendszerek jelzetei. Az első három az ETO és a TO mezője; a 084-es lehet bármilyen hierarchikus vagy enumeratív rendszer, a 085 pedig a klasszikus szintetizáló osztályozási rendszer mezője.

Magyarországon a könyvtárak többsége az ETO-t, néhány a TO-t használja; a másik két mező értékei valószínűleg nem saját feldolgozásból erednek (egyes ETO-t használó könyvtárakban a TO értékei se), hanem a rekordok átvételével keletkezettek (8. táblázat).

8. táblázat

A 0XX mezőcsoport tartalmi azonosítói

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
080 ETO jelzet										
\$???	-	-	-	-	2	0,00	-	-	-	-
\$# ???	-	-	-	-	-	-	-	-	19	0,00
\$! ???	-	-	-	-	1	0,00	-	-	-	-
\$0 bes, rekordazon, és szabv??	-	-	-	-	6	0,00	58	0,00	122	0,00
\$1 URI	-	-	-	-	1	0,00	28	0,00	89	0,00

TMT 67. évf. 2020. 11.sz.

\$2 kiadásazonosító	-	-	-	-	3	0,00	127	0,01	607	0,01
\$3 ???	-	-	-	-	1	0,00	126	0,01	453	0,01
\$4 ???	-	-	-	-	1	0,00	-	-	13	0,00
\$5 ???	-	-	-	-	-	-	49	0,00	4362	0,05
\$6 kapcsolódás	-	-	-	-	1	0,00	93	0,01	269	0,00
\$7 ???	-	-	-	-	-	-	-	-	157	0,00
\$8 mezőkapcs, és sorsz,??	-	-	2	0,00	16	0,00	126	0,01	549	0,01
\$9 ???	-	-	-	-	13	0,00	53	0,00	62	0,00
\$A	-	-	-	-	4	0,00	-	-	12	0,00
\$a ETO jelzet	-	-	100 245	5,68	535 691	50,91	1 038 987	81,85	3 566 807	40,12
\$b Cutter példányazon	-	-	5	0,00	3	0,00	14	0,00	1202	0,01
\$c ???	-	-	-	-	-	-	8	0,00	458	0,01
\$d ???	-	-	-	-	-	-	40	0,00	113	0,00
\$e ???	-	-	-	-	-	-	1	0,00	3	0,00
\$g ???	-	-	-	-	-	-	12	0,00	17	0,00
\$h ???	-	-	-	-	-	-	-	-	10	0,00
\$i ???	-	-	-	-	1	0,00	-	-	1	0,00
\$j ???	-	-	-	-	-	-	-	-	1	0,00
\$l ???	-	-	-	-	-	-	-	-	5	0,00
\$m ???	-	-	-	-	-	-	2	0,00	-	-
\$n ???	-	-	-	-	-	-	2	0,00	5	0,00
\$p ???	-	-	-	-	-	-	2	0,00	3	0,00
\$q???	-	-	-	-	-	-	2	0,00	5	0,00
\$r ???	-	-	-	-	-	-	2	0,00	1	0,00
\$s ???	-	-	-	-	-	-	16	0,00	53	0,00
\$t ???	-	-	-	-	-	-	-	7	7	0,00
\$v ???	-	-	-	-	-	-	-	-	22	0,00
\$w ???	-	-	-	-	-	-	1	0,00	-	0,00
\$x általánosan közös alosztás	-	-	26	0,00	6	0,00	81	0,01	184	0,00
\$y	-	-	-	-	-	-	-	-	4	0,00
\$z	-	-	-	-	-	-	-	-	2	0,00
082 TO										
\$0,	-	-	-	-	-	-	-	-	2	0,00
\$2 kiadásazon??,	7 365 077	38,70	112 572	6,38	-	-	50	0,00	37 461	0,42
\$3 ???	-	-	-	-	-	-	-	-	2	0,00
\$5 ???	-	-	-	-	-	-	-	-	1	0,00
\$7 ???	-	-	-	-	-	-	-	-	1	0,00
\$8 mezőkapcs, és sorsz,	-	-	264	0,01	-	-	-	-	4	0,00
\$a TO jelzet	7 365 118	38,71	192 905	10,39	1	0,00	745	0,06	82 915	0,93
\$b Cutter példányazon,	-	-	4 082	0,23	-	-	1	0,00	74	0,00
\$d ???	-	-	-	-	-	-	-	-	7	0,00
\$q jelzetkiadó intézmény	-	-	493	0,03	-	-	-	-	8	0,00
\$v ???	-	-	-	-	1	0,00	-	-	-	-
\$x ???	-	-	769	0,04	-	-	-	-	-	-

Ungváry R.: Ismeretszervező-könyvtári tartalmi feltárásának összehasonlító ...

083a további TO jelzet	–	–	582	0,03	–	–	7	0,00	21	0,00
084\$a egyéb oszt, jelzet	4 602 135	24,18	52 588	2,98	1	0,00	5 136	0,40	87	0,00
085\$a szintetizált oszt, jelzet	1 348 402	7,09	–	–	–	–	3	0,00	8	0,00

A DNB-ben csak a TO használatos; az állomány rekordjainak jelentős része tartalmazza ezt a jelzést. A DNB gyakorlata reálisan csak az OSZK gyakorlatával volna összehasonlítható. Az egyéb és a szintetizált osztályozási rendszer használata is a DNB-re (és a Finn Nemzeti Könyvtárra) jellemző.

Egészében úgy tűnik, hogy a külföldi könyvtárakban valamivel több osztályozási rendszert alkalmaznak, mint Magyarországon.

Mindez arra utal, hogy a könyvtári világban *továbbra is kitüntetett fontossága van* ezeknek a klasszikus, mesterséges nyelven alapuló osztályozási rendszereknek. Az igazi az volna, ha adott tárgyszóval keresve opcionálisan meg lehetne adni, hogy ha az adott tárgyszó az ETO, vagy valamilyen más rendszer jelzetének mutatószavával egyezik, akkor opcionálisan kérhetők, hogy a találatokban az ezzel a jelzettel osztályozott források is bekerüljenek. Kétségtelen, hogy ez egyelőre vágyálom, noha technikailag semmiféle nehézséget nem jelent a megvalósítás.¹⁵

A magyar könyvtárakban az ETO alkalmazásának kultúrája közel évszázados hagyományra tekinthet vissza, és ezzel magyarázható, hogy a használata a nemzeti és a tudományos könyvtárakban is egyaránt jelentős.

A SZTE hibás ETO-almezői és a nem használt TO-almezők letöltések következtében keletkeztek, javításuk időközben a SZTE-be megtörtént. Király Péter felmérése adott a felismerésükre lehetőséget – a könyvtári szoftverek nem.

A MTAK 080 mezőjében a \$0, \$1, \$b és \$x az értékek vagy egyáltalán nem felelnek meg az almezők tartalmának, vagy sok nem odavaló értéket is tartalmaznak. Az ismeretlen almezők pe-

dig túlnyomórészt valamiféle jelzeteket tartalmaznak.

A SZTE 080 mezőjében hasonlókat tapasztalhatók. Az ismeretlen almezők ugyancsak túlnyomórészt valamiféle jelzeteket tartalmaznak. A 080 mező HUNMARC \$c és \$d almezői többnyire ugyancsak nem felelnek meg a HUNMARC szerinti almezők tartalmának. A SZTE 090 lokális mezőjében a legkülönösebb, hogy két két \$a és \$b almező van, különböző értékekkel.

Érdekes, hogy az adatokat szemügyre véve a SZTE 084-es adatainak jelentős része nem került be a MOKKA-ba. Az MTAK \$! almezőjének az értéke sem került be a MOKKA-ba, más könyvtárból vagy könyvtárakból viszont a \$# almező értékei bekerülhettek. Nyilván összefügghet ez a betöltéskor használt ellenőrző program minőségével, melyről nem tudni semmit.

Mindez relatíve nem sok érték, de azért egészében mégis kiábrándító, és ahogy utaltam rá, nem elsősorban a könyvtári feldolgozókon múlik, hanem, a katalóguskezelő szoftver elégtelen ellenőrző képességein.

4. A 6XX tárgyszavak mezőcsoport

4.1 A 600-as személynév tárgyszó

A tulajdonnév tárgyszavak mezői közül (a különleges esetet képviselő 651 földrajzi nevekén kívül) csak a 600-as, 651-es, 653-as és 655-ös mezők összehasonlítását mutatom be. A 610-es, 611-es és 630-as mező az előbbiekhöz hasonló jellegzetességekkel rendelkezik, a 647 esemény és 654 fazettás tárgyszó mezőt pedig nem használják (9. táblázat).

9. táblázat

A 600-as mező összehasonlítása

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
600 személynév tárgyszó										
\$# ???	–	–	–	–	–	–	–	–	193	0,00
\$0 besorolási rekordazon,	378 007	1,99	3 860	0,22	3	0,00	1	0,00	12	0,00
\$1 URI	–	–	1	0,00	–	–	–	–	6	0,00
\$2 forrás	433 870	2,27	79 207	4,49	1	0,00	63 060	4,97	73 665	0,83
\$3 dokumentumrész	–	–	1	0,00	–	–	81	0,01	479	0,01
\$4 kapcsolódó URI	–	–	49	0,00	3	0,00	–	–	177	0,00
\$6 kapcsolódás	–	–	280	0,00	164	0,01	–	–	19	0,00
\$8 mezőkapcsoló és sorsz,	18 733	0,10	–	–	1	0,00	–	–	4	0,00
\$D ???	–	–	–	–	–	–	–	–	1	0,00
\$M ???	–	–	–	–	–	–	–	–	1	0,00
\$a rendszó (személynév)	433 676	2,28	79 207	4,49	172 001	16,35	64 228	5,06	490 403	5,52
\$b uralkodói sorszám	88 640	0,05	4 343	0,25	5 582	0,53	3 788	0,30	27 158	0,31
\$c méltóság, foglalkozás	39 799	0,21	16 501	0,94	29 909	2,84	8 245	0,65	125 575	1,41
\$d kronologikus kiegészítő	351 243	1,85	61 884	3,51	139 159	13,22	6 592	0,52	391 773	4,41
\$e kapcsolatjelölő	–	–	67	0,00	23 725	2,25	2	0,00	272	0,00
\$f a mű keletkezési ideje	82	0,00	15	0,00	2	0,00	–	–	47	–
\$g állandósult melléknév	292	0,00	80	0,00	722	0,07	3	0,00	1429	0,02
\$h hordozó	–	–	–	–	1	0,00	–	–	–	–
\$i ???	–	–	–	–	–	–	–	–	3	0,00
\$j egyéni név egyéb elem	–	–	–	–	6	0,00	–	–	2 313	0,03
\$k formai alcím ¹⁶	–	–	36	0,00	–	–	–	–	3	0,00
\$l mű nyelve	–	–	61	0,00	1	0,00	–	–	28	0,00
\$m zenei változat hordozója ¹⁷	212	0,00	15	0,00	238	0,02	–	–	416	0,00
\$n mű részének száma	819	0,00	186	0,01	4	0,00	1	0,00	16	0,00
\$o zenei átdolgozás adata	–	–	–	–	11	0,00	–	–	–	–
\$p mű részének címe	478	0,00	147	0,01	–	–	–	–	16	0,00
\$q név teljesebb formája	–	–	2 297	0,13	3	0,00	–	–	31	0,00
\$R ???	–	–	–	–	–	–	–	–	3	0,00
\$s változat	21	0,00	6	0,00	2	0,00	2	0,00	18	0,00
\$t címreláció	36 496	0,19	9 398	0,53	10 790	1,03	50	0,00	13 451	0,15
\$' ???	–	–	–	–	–	–	–	–	1	0,00

A MOKKA nagy betűs hívójelű almezőiben az értékek értelmezhetetlen zajt képviselnek.

A viszonylag új \$0 *besorolási rekordazonosító* almező a magyar katalógusokban még alig fordul elő. A DNB-ben viszont a személynevek nagy része már rendelkezik besorolási rekordazonosítóval, és bekerül a közös egységesített németországi tárgyszójegyzékbe (Gemeinsame Normdatei), miáltal a 600-as mezőbe kerülő nevek többségé-

nek ez lett a forrása. Hasonló lehet a helyzet a GENT-ben is. A SZTE-ben ugyancsak forrásként kezelik az általuk feldolgozott egységesített személyneveket.

A szintén új \$1 *URI* mezőt (*real world object URI*) pedig még sehol sem használják. Megjegyzendő, hogy mindkét almező viszonylag új fejlemény a MARC21-ben, nem is biztos, hogy a vizsgált könyvtárak katalóguskezelő rendszereiben újként

definiálhatók, ha a rendszert úgy telepítették, hogy abba minden eleve „bele lett égetve”, és a felhasználói részről nem lehet rajtuk módosítani.

A többi 6XX-es tárgyszómezőre is ez érvényes. A \$0 almezőt még a leginkább a DNB-ben használják, de ez a gyakorlat csak az OSZK-val lenne összehasonlítható.

Különös jelenség, hogy a DNB \$2 forrás almezőkben számszerűen egész kicsit több a forrásértékek száma, mint a \$a rendszó almezőben a rendszóértékeké. Megint csak valószínűsíthető, hogy ebben elsősorban a szoftver a ludas, amely ezt jelzés nélkül megengedi. A személynevek felvétele általában nem személynév-szótárakból, hanem a források alapján történik, ezért a forrásértékek száma általában kisebb, mint a rendszóértékek száma.

Az MTAK katalógusában használják viszonylag a leginkább a személynév tárgyszót (\$a rendszó 16,35%). Ennek a könyvtárnak néhány más almezőjében is (\$c uralkodói sorszám és \$d méltó-

ság) az átlagosnál több az érték. A SZTE katalógusában a \$2 forrás almező kitöltöttsége majdnem összhangban van egymással (a vonatkozó almezők értékeinek teljes azonossága érdekében Király Péter felmérésének ismeretében folyamatban van). A magyar könyvtárakra a leginkább jellemző a \$c és \$d használata.

A \$d kronologikus kiegészítő használatát a SZTE könyvtárban hasonló nevek esetében nem tekintik kielégítőnek; a személyek besorolási adatait igyekeznek a VIAF azonosítóval pontosítani. Hasznos lenne tudni, hogy miért, és hogy más könyvtárakban is felmerült-e ilyen probléma a \$d almezővel kapcsolatban.

4.2 A 648-as kronologikus tárgyszó

A kronologikus (évszám, évtized, évszázad, évezred, stb.) tárgyszavak elkülönített minősítése szoros összefügg a 045 rekord tartalmának időtartománya és 046 különlegesen kódolt dátumok mezőkkel és azok összehangolt használatával (10. táblázat).

10. táblázat

A 648-as mező összehasonlítása

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
648 kronologikus tárgyszó										
\$0 besorolási rekordazon,	–	–	147	0,01	–	–	–	–	–	–
\$2 forrás	439379	2,31	4976	0,28	–	–	–	–	3	0,00
\$a rendszó (kronologikus)	439379	2,31	4976	0,28	–	–	–	–	519	0,01

A DNB-ben és az összehasonlításba nem bevont Finn Nemzeti Könyvtárban használják nagyjából egyforma arányban ezt a mezőt, és a \$2 meg a \$a almezők kitöltöttsége egyforma. A külföldi tudományos könyvtárakra a használat lényegében nem jellemző, a magyar tudományos könyvtárakra még kevésbé. A MOKKA-ban megjelenő értékek történelmi szakkönyvtárból származnak. Mivel a Köztársaságban a kronologikus tárgyszavak eleve minősítve vannak a 148-as mezőben az OSZK katalógusában eleve kényszerűen a 648-as mezőbe kerülnek a kronologikus tárgyszavak.

A csekély használatnak talán az is oka lehet, hogy a katalogizáló szoftver nem minősíti (vagy nem minősítik benne) külön a kronologikus tárgyszavakat, hanem hozzácsapja a szaktárgyszavakhoz.

A Finn Nemzeti könyvtár kivételével, ahol a 648\$a kronologikus tárgyszó mellett a 045 rekord tartalmának időtartománya (kora, korszaka) mező használata is relatíve jelentős (4.9%), a többi vizsgált katalógusból ez a mező, ugyanúgy mint a 046 esemény, tartalom dátuma mezők teljesen hiányoznak. Pedig ezek a tartalmi keresés szempontjából nagyon fontosak lennének.

A 648 időtárgyszavak és a 647 esemény tárgyszavak elkülönített kezelésének a hiánya, továbbá a 045 rekord tartalmának időtartománya nagyon ritka használatának oka nem elsősorban erőforrások hiányára vezethető vissza. Lehet, hogy egyelőre hiányzik a szakmai érzékenység ahhoz, hogy az időtárgyszavaknak jelentőséget tulajdonítsanak. Erre utalhat az is, hogy csak a GENT-ben fordulnak elő besorolási rekordazonosítóval ellátott kronologikus tárgyszavak, a többi vizsgált katalógusban nem.

A 647 esemény tárgyszó használatának teljes hiánya is arról tanúskodik, hogy a tárgyszavak idő szerinti tipologizálásának nem tulajdonítanak külö-

nösebb szerepet. A találatok tárgyszótípusonkénti elkülönítésének (személy a keresések értékelését támogató szerepe) idővel, amikor már az automatikus osztályozás és mesterséges intelligencia-rendszerek kinövik gyermekcipőiket, további szerepük is lesz.

4.3 A 650-es szaktárgyszó

A mező a szakterületi tartalmi feltárás központi eleme. Sokat elárul a tartalmi feldolgozás minőségéről. Ahogy az ebben a tanulmányban nem érintett 1XX-es mezők a bibliográfiai-formai feltáráról (11. táblázat).

11. táblázat

A 650-es mező összehasonlítása

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
650 szaktárgyszó										
\$# ???	-	-	-	-	-	-	-	-	7 118	0,08
\$\$???	-	-	-	-	-	-	-	-	74	0,00
\$0 besorolási rekordazon,	2 844 292	14,95	67 743	3,84	-	-	-	-	72	0,00
\$1 URI	-	-	-	-	-	-	-	-	53	0,00
\$2 forrás	3 612 054	18,98	689 532	39,08	64	0,01	733 857	57,82	1 402 485	15,78
\$3 dokumentumrész	-	-	-	-	-	-	1 922	0,15	8848	0,10
\$5 kapcsolódó intézm, ???	-	-	-	-	-	-	-	-	2	0,00
\$4 kapcsolat	-	-	-	-	-	-	-	-	108	0,00
\$5	-	-	-	-	-	-	-	-	10	0,00
\$6 kapcsolódás	-	-	337	0,02	-	-	-	-	2	0,00
\$7 ???-	-	-	-	-	-	-	-	-	2	0,00
\$8 mezőkapcs, és sorsz,	209 766	1,10	2	0,00	-	-	-	-	-	-
\$A ???	-	-	-	-	-	-	-	-	1	0,00
\$B ???	-	-	-	-	-	-	-	-	2	0,00
\$H ???	-	-	-	-	-	-	-	-	1	0,00
\$M ???	-	-	-	-	-	-	-	-	1	0,00
\$P ???	-	-	-	-	-	-	-	-	1	0,00
\$S ???	-	-	-	-	-	-	-	-	5	0,00
\$T ???	-	-	-	-	-	-	-	-	6	0,00
\$V ???	-	-	-	-	-	-	-	-	2	0,00
\$X ???	-	-	-	-	-	-	-	-	1	0,00
\$Z ???	-	-	-	-	-	-	-	-	3	0,00
\$a rendszó (tárgyszó)	3 612 054	18,98	689 532	39,08	140 030	13,31	733 809	57,82	3 119 811	35,10
\$b földr, nevet köv, tárgyszó	-	-	50	0,00	12	0,00	26	0,00	13 808	0,16

Ungváry R.: Ismeretszervező-könyvtári tartalmi feltárásának összehasonlító ...

\$c esemény helye	-	-	2 690	0,15	13	0,00	17	0,00	1 716	0,02
\$d esemény időtartama	-	-	3 363	0,19	126	0,00	35	0,00	3 261	0,04
\$e kapcsolatjelző	-	-	9	0,00	4	0,00	3	0,00	22	0,00
\$f ???	-	-	-	-	3	-	7	0,00	33	0,00
\$g állandósult melléknév	208 211	1,09	29	0,00	1	0,00	22	0,00	55	0,00
\$h ???	-	-	-	-	-	-	-	-	10	0,00
\$i ???	-	-	-	-	-	-	2	0,00	40	0,00
\$j ???	-	-	-	-	1	0,00	1	0,00	12	0,00
\$k ???	-	-	-	-	-	-	9	0,00	55	0,00
\$l ???	-	-	-	-	-	-	5	0,00	43	0,00
\$m ???	-	-	-	-	-	-	1	0,00	30	0,00
\$n ???	-	-	-	-	-	-	3	0,00	30	0,00
\$o ???	-	-	-	-	1	0,00	2	0,00	4	0,00
\$p ???	-	3	0,00	-	-	-	6	0,00	48	0,00
\$r ???	-	-	-	-	-	-	6	0,00	15	0,00
\$s ???	-	3	0,00	-	-	-	5	0,00	41	0,00
\$t ???	-	1	0,00	-	-	-	26	0,00	102	0,00
\$u ???	-	-	-	-	-	-	-	-	5	0,00
\$' ???	-	-	-	-	-	-	-	-	¹⁸	0,00
\$v formai altárgyszó	-	-	63 791	3,62	43 158	4,10	10	0,00	64 029	0,72
\$x általános altárgyszó	334 101	1,76	312 975	17,74	120 712	11,47	622 101	49,01	1 447 133	16,28
\$y kronologikus altárgyszó	-	-	61 933	3,51	70 134	6,66	170 494	13,43	501 504	5,64
\$z földrajzi altárgyszó	-	-	173 050	9,81	65 392	6,21	669	0,05	369 933	4,16

A DBN-ben a \$0 *besorolási rekordazonosítónak* gyakrabban van értéke (14.95%), mint a 600-as mező esetén, és ez a helyzet a \$a *rendszo* esetén is, melynek értékei teljesen összhangban vannak a \$2 *forrás* almező értékeivel. Ez bizonyára a besorolási tárgyszórendszerének (Deutsche Normdatei) a gondozottságával is összefügg. Viszont a többi \$n–\$e almező teljesen kitöltetlen. (Az OSZK-val való összehasonlítás lehetősége persze itt is hiányzik.)

A \$2 *forrás* és \$a *rendszo* almezők értékei között az MTAK kivételében teljes az összhang, így a DNB mellett a SZTEK katalógusában is. Tapasztalatom szerint az OSZK-ban ennek összhangját a rendszer eleve biztosítja, mivel minden szaktárgyszó kötelezően a Köztauruszból, és földrajzi név a Köztaurusz Geotaurusz részéből származik. Ezt az egyetemes tezaurszt számos más magyar könyvtárban is használják.

Figyelemre méltó, hogy a SZTE katalógusában mind a 080\$a *ETO-jelzet* almező, mind a 650\$a *szaktárgyszó* almező értékeinek száma viszonylag nagy (81.85%, ill. 57.82%). Mindez gondos tartalmi feltárára utal. Ez figyelhető meg a Finn Nemzeti Könyvtár, és tapasztalatom szerint ez az OSZK esetében is.

A SZTE-ben \$x *általános altárgyszót*, és a 648-as mező helyett a \$y *kronologikus altárgyszót* használják. A \$x *formai altárgyszót* viszont alig, a \$z *földrajzi altárgyszónak* pedig vannak értékei, holott a 651 földrajzi tárgyszót is használják. Lehet, hogy a hagyomány is szerepet játszik ebben?

A sok értelmezhetetlen – valószínűsíthetően hibás – almező kiábrándító. Nem a figyelmetlenség miatt, mert a hibák mennyisége hibahatáron belüli, hanem mert fényt vetnek a könyvtári szoftverek ellenőrző képességeinek a gyengeségére.

4.4 A 651-es földrajzi tárgyszó

A földrajzi tárgyszó a keresés szempontjából gyakran szervesen kiegészíti a szaktárgyszavak által megnevezett fogalmakat (12. táblázat).

12. táblázat

A 650-es mező összehasonlítása

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
651 tárgyszó										
\$0 besorolási rekordazon,	984 088	5,17	12 199	0,69	–	–	–	–	2	0,00
\$1 URI	–	–	1	0,00	–	–	–	–	8	0,00
\$2 forrás	1 114 929	5,86	84056	4,76	1	0,00	35 122	2,77	66 400	0,75
\$3 dokumentumrész	–	–	–	–	–	–	113	0,01	1 542	0,02
\$4 kapcsolódó URI	–	–	–	–	–	–	–	–	60	0,00
\$5 kapcsolódó intézm, ???	–	–	–	–	–	–	–	–	3	0,00
\$6 kapcsolódás	–	–	57	0,00	–	–	–	–	1	0,00
\$7 ???	–	–	–	–	–	–	–	–	1	0,00
\$8 mezőkapcs, és sorsz,	41 490	0,22	–	–	–	–	–	–	6	0,00
\$9 ???	728	0,00	–	–	–	–	–	–	–	–
\$@ ???	–	–	–	–	–	–	–	–	1	0,00
\$M ???	–	–	–	–	–	–	–	–	2	0,00
\$a rendszó (tárgyszó)	1 114 929	5,86	84 056	4,76	7 325	0,70	50 893	4,01	674 195	7,58
\$b tárgyszót követ, földr, név	–	–	4	0,00	4	0,00	8	0,00	3 495	0,04
\$c kiegészítő elem	–	–	–	–	6	0,00	2	0,00	10 042	0,11
\$d ???	–	–	–	–	2	0,00	2	0,00	15	0,00
\$e kapcsolatjelölő	–	–	–	–	4	0,00	–	–	1	0,00
\$f ???	–	–	–	–	–	–	1	0,00	1	0,00
\$g állandósult melléknév	78 322	0,41	19	0,00	124	0,01	–	–	1	0,00
\$h ???	–	–	–	–	–	–	–	–	1	0,00
\$i	–	–	–	–	–	–	1	0,00	2	0,00
\$k	–	–	1	0,00	–	–	1	0,00	8	0,00
\$l	–	–	–	–	–	–	–	–	2	0,00
\$m ???	–	–	–	–	–	–	3	0,00	5	0,00
\$n ???	–	–	–	–	1	0,00	–	–	–	–
\$p ???	–	–	–	–	–	–	–	–	1	0,00
\$s ???	–	–	–	–	–	–	–	–	2	0,00
\$t ???	–	–	–	–	–	–	–	–	23	0,00

A 650-es mezőről mondottak nagyjából a 651-es mező esetében is érvényesek. A \$2 forrás és a \$a földrajzi név almezők között azonban sem az MTAK, sem a SZTE esetében nincs összhang. (Az OSZK-ra vonatkozóan a Köztaurusz használata következtében a 648-as, 650-es és 651-es me-

zőkre mondottak érvényesek: mindegyik tárgyszavai kötelezően a Köztauruszból származnak.)

A HUMARC szerint a \$c kiegészítő adat (hátravetett értelmező) almező a homonim tárgyszavak és homonim földrajzi nevek megkülönböztetésére (ez

a MARC21-ben nem foglalt) való. Az almezőt az OSZK-ban használják.

4.5 A 653-as szabadon választható tárgyszó

Ez a tárgyszó szabadon, az adott könyvtár katalógusában használt kötött szótár igénybevétele nélkül, ezektől teljesen függetlenül vehető föl (nevezik „indexkifejezésnek” is, de mutatószó bármelyik 6XX-es mező tárgyszava lehet). Más szóval eleve nem is lehet olyan tárgyszó, mely egy ilyen kötött szótárból származnék. Például az OSZK-ban a

Köztaurusz tárgyszava nem kerülhet a 653\$a almezőjébe.²⁵ A gyakorlatban természetesen előfordul, hogy a feldolgozó valamilyen kézikönyv, netán akár tárgyszórendszer alapján veszi föl ezt a tárgyszót, de az adott forrás használata rendkívül véletlenszerű, a forrásnak nincs igazán meghatározó jelentősége (13. táblázat).

Ezért ennek a mezőnek csak egyetlen szabványos almezője van, a \$a *rendszo*.

13. táblázat

A 653-as mező összehasonlítása.

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
653 szabadon választott tárgyszó										
\$0 besorolási rekordaz ??? ²¹	–	–	–	–	–	–	4	0,00	2	0,00
\$2 forrás ???	–	–	–	–	–	–	4	0,00	4	0,00
\$3 kapcsolódás	–	–	–	–	1	0,00	–	–	15	0,00
\$6 mezőkapcsolat és sorszám	–	–	–	–	307	0,03	–	–	–	–
\$A ???	–	–	–	–	1	0,00	–	–	1	0,00
\$a rendszer (szab, v, tárgyszó)	6 116 045	32,14	–	–	68 550	6,51	122	0,01	327 527	3,68
\$c ???	–	–	–	–	1	0,00	–	–	1	0,00
\$d ???	–	–	–	–	1	0,00	–	–	2	0,00
\$j ???	–	–	–	–	3	0,00	–	–	–	–
\$l ???	–	–	–	–	1	0,00	–	–	2	0,00
\$m ???	–	–	–	–	1	0,00	–	–	1	0,00
\$n ???	–	–	–	–	1	0,00	–	–	1	0,00
\$t ???	–	–	–	–	1	0,00	–	–	1	0,00
\$x ???	–	–	–	–	1	0,00	–	–	320	0,00
\$y ???	–	–	–	–	1	0,00	–	–	2	0,00
\$z ???	–	–	–	–	1	0,00	–	–	4	0,00

A fent mondottak fényében érthetetlen, hogyan lehet egy ilyen tárgyszónak akár besorolási rekordazonosítója (hiszen nem besorolási adat), forrásadata stb. Nem beszélve a többi érthetetlen almezőről. Ha egy tárgyszót bármilyen idegen, más intézmény által gondozott tárgyszórendszerből veszünk föl, akkor azt a 650, 651 illetve 655 mezőben kell rögzíteni, megadva a \$2 *forrás* almezőben az adott forrást. Az ilyen tárgyszavaknak természetesen lehet \$0 *besorolási rekordazonosítójuk* is.

4.6 A 655-ös formai tárgyszó

A 655-ös mezőbe kerülnek a dokumentum műfaját, típusát, formáját megnevező tárgyszavak. Mint ilyenek, a 008-as mező értékeihez hasonlóan fontos – ha nem még fontosabb – kiegészítői lehetnek a tartalom szerinti keresésnek. Vannak kivételek, köztük egy jelentős: a Kongresszusi Könyvtárban hosszú ideig nem különböztették meg a formai tárgyszavakat a szaktárgyszavaktól – talán éppen azon az alapon, hogy fontosságuk a szaktárgyszavakéval egyenrangú (14. táblázat).

14. táblázat

A 655-ös mező összehasonlítása

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
655 formai tárgyszó										
\$0 besorolási rekordazon,	5 087 226	16,74	15 908	0,85	–	–	–	–	6	0,00
\$2 forrás	5 098 459	26,80	197 768	11,21	62	0,01	1 493	0,12	260 013	2,92
\$3 dokumentumrész	–	–	–	–	–	–	–	–	1	0,00
\$4 ???	–	–	–	–	1	0,00	–	–	1	0,00
\$a rendszó (formai tárgyszó)	5 098 603	26,80	197 768	11,21	19 226	1,83	1 493	0,12	465 447	5,24
\$b nem fókusz kifejezés	–	–	3	0,00	–	–	–	–	–	–
\$c fazettás fölérendelt	–	–	4	0,00	–	–	–	–	–	–
\$i ???	–	–	–	–	–	–	–	–	4	0,00
\$s ???	–	–	–	–	–	–	–	–	1	0,00
\$t ???	–	–	–	–	–	–	–	–	1	0,00
\$t ???	–	–	–	–	–	–	–	–	4	0,00

A mező alkalmazása kiegyensúlyozottabb képet mutat; de ahogy a többi 6XX mező \$2 forrás almezője esetében, az MTAK katalógusában a 655-ös mezőben sincs összhang a \$a rendszó almező és a \$2 forrás értékei között. Talán nem tartják annyira fontosnak a forrásadatot?

A SZTE katalógusában a 655 formai tárgyszó mezőt alig használják. Ebben a mezőben a \$v formai altárgyszó egyáltalán nem fordul elő, a \$x általános altárgyszónak is csak kevés értéke van.

5. Összegzés

5.1 Feldolgozási jellegzetességek

A mezők és almezők felhasználása

Vannak mezők, mint például a 045 rekord tartalmának időtartományja és 046 különlegesen kódolt dátumok, továbbá a 648 kronologikus tárgyszó, melyeket vagy egyáltalán nem, vagy csak elenyésző mértékben használnak. Érdekes ezt a gyakorlatot felülvizsgálni.

A 653 szabadon választott tárgyszó mezőben megjelenő hibás almezők arra utalhatnak, hogy ezeknek a tárgyszavaknak az alkalmazásához egyelőre félreértések is kapcsolódnak. Mintha elhomályosulna, hogy ezek a fajta tárgyszavak elvileg nem forrásból származnak, mert ha igen, akkor a többi tárgyszómezőben a helyük, ahol a forrás is megadható.

A \$0 besorolási rekordazonosító és a \$1 URI almezők használata idővel egyre fontosabb lehet.

Számos almező használata, illetve nem használata az egyes katalógusokban meglehetősen különböző. Ennek okait az adott könyvtárak szakemberei ismerik a legjobban.

Hibás almezők

Keletkezésük okait a könyvtári rendszerek minőségével foglalkozó fejezetben tárgyalom. Itt csak annyit említek meg erről, hogy az MTAK-ban rendkívül gondos munkát végeztek az HUNMARC-ról a MARC21-re való átállítás érdekében, és erről impozáns teljességű dokumentum is született [2]. Bizonyára értik a könyvtár szakemberei, hogy ennek ellenére miért található mégis hibás almezőhívójelek és hibás 008-as mezőértékek Király Péter e tanulmányhoz alapul vett elemzésének MTAK-táblázataiban.

A mezők és almezők kitöltöttsége

Egyes katalógusokban a mezők és almezők kitöltöttségének mértéke feltűnően egyenlőtlen. Ezek a különbségek elválaszthatatlanok az országok, és azokon belül az egyes könyvtárak kialakult gyakorlatától, szemléletétől. Ezek még az olyan, látszólag pontosan formalizált esetekben is, mint az egységesített személynevek megformálásának gyakorlata, meglehetősen nagy eltéréseket eredményezhetnek, amihez elég a VIAF egyelőre csak személyneveket tartalmazó rendszerébe bepillantani, és mondjuk megnézni, hogyan fest az egyes nemzeti rendszerekben Krúdy Gyula egységesített névalakja (4. ábra).



4. ábra Krúdy Gyula egységesített neve a VIAF-ban
(Virtuális Nemzetközi Besorolási Állomány, <http://viaf.org/>)

A kérdés valójában szorosan kapcsolódik az egyes nemzeti nyelvek jellegzetességeihez is. Ez a fogalmak szaktárgyszavakkal azonosított névválasztásában különösen megnyilvánul. A nevek tartalma nyelvenként nem teljesen azonos, ezért a jövő többnyelvű keresőrendszereiben komoly feladat lesz az összehangolásuk a többnyelvű névterekben.

Bizonyos szempontból összetartozó almezők közötti koherencia

Hibáinak okait a könyvtári rendszerek minőségével foglalkozó fejezetben tárgyalom.

A lokális almezők alkalmazása

Feltűnő, hogy milyen kevés a számuk. Lokális \$9 almező ráadásul eleve mindig csak egyetlen lehetne egy mezőn belül, holott ha az igény rá felszínre kerülne, sokszor egy mezőn belül is több kellhet ezekből.

Se a MARC21-ben, se a HUNMARC-ban nem szereplő mezők és almezők

Ilyenekből viszonylag sok van. Jelentős részük valószínűleg olyan hiba következménye, mely megfelelő ellenőrzési lehetőségekkel rendelkező, jól felkészített könyvtári szoftver esetében nem következnek be. Az is lehet azonban, hogy kisebb részük valójában lokális almező. Az említett, egyelőre még rejtett igény kielégítése a MARC21 további fejlesztésének a feladata.

5.2 A könyvtári rendszer minősége

A legtöbb és a legnagyobb problémát ezek a szoftverek okozzák. Az adatelemek egy része, a mezőhívójelek és almezőjelek, valamint az előírt egykarakteres indikátorok formálisan viszonylag teljes körűen ellenőrizhetők.

A lokális mezőkhöz és almezőkhöz kimerítőbb tájékoztatás is definiálható lenne egy kezelőrendszerben.

A szoftvereknek olyanoknak kell lenniük, hogy új mezőket, almezőket, indikátorokat, továbbá jogosultságokat maguk az erre jogosított nem informatikus feldolgozók tudjanak fölvenni.

A speciális 008 és 0XX mezők közül a tartalommal összefüggőknek a bevonása a 6XX mezőcsoport tárgyszavaival végzett keresésbe jelenleg túlnyomórészt nincs a laikus végfelhasználóknak megfelelően megoldva. Beleértve a 008-as mező dokumentumtípus-értékeivel kombinálható keresést. Ugyanez vonatkozik mesterséges nyelven alapuló osztályozási rendszerek (ETO, TO stb.) jelzeteire, melyek szerinti keresés sincs összehangolva a természetes nyelven alapuló tárgyszavakkal végzett kereséssel. Azaz hiányzik az ETO, TO stb. mutatószavainak a tárgyszavakkal való összehangolása.

A kezelőrendszerek egy része nincs felkészítve arra, hogy egymással összefüggő mezők/almezők koherenciáját kellőképpen ellenőrizzék (például a 008 mező 35–37. nyelv pozíciójának és a 040\$b nyelv almezőnek az értékei között).

A felsorolás korántsem teljes.

5.3 A mintaként kiválasztott katalógusok összehasonlítása

Megfigyelhető, hogy a külföldi könyvtárakban a hibás almezők száma kisebb. Ennek elsősorban nem a magyar szakemberek felkészültsége az oka, hanem az előző pontban tárgyalt könyvtári szoftver elégtelenségei – és nem utolsósorban az anyagi ellátottság.

6. Kitekintés a MARC21-re általában

Jason Thomale [13] és utána Karen Coyle [3] a MARC21 iránti megértő, ugyanakkor kritikus, és mélyre ható elemzésükben kimutatták, hogy a MARC21-formátumban van néhány olyan vonás, amely miatt az automatizált hasznosítás esetén problémák merülnek föl. Arra a tárgyalási részletességre és mélységre, mely a két közleményt jellemzi, itt nem térhetek ki. Egyet említek csak, a legfeltűnőbb atavizmust, mely valójában ma már teljesen felesleges: fontos adatmezőkben, például a főcímben intellektuálisan rögzítenek központoszási jeleket.

Például a kiadási adatok esetén:

300\$a 303 pages : \$b illustrations (chiefly color)
; \$c 26 cm

Például az esemény (a Boston közelében levő Bunker Hilli csata) tárgyszavak esetén:

6477\$a Bunker Hill, Battle of \$c (Boston, Massachusetts : \$d1775) \$2 fast \$0 (OCoLC)
fst01710024

Az ISBD formátumú megjelenítés esetén természetesen a MARC-mezők és -almezők által meghatározott központoszási jeleket alkalmazzák, ezért ezek átformálása önálló almezőkre formálisan biztosítva van. A rekordfej 18. a *bibliográfiai leírás szabálya* pozíciójának egy karakteres értékével és a 040\$e almezőben szövegesen adható meg a katalógizálási szabályzat, amely alapján elvileg tudható, kell-e az adott, konvertálandó állományban központoszási jelekre számítani. Az más kérdés, hogy ezt mennyire gondosan töltik ki a feldolgozók. A SZTEK például használ központoszási jeleket.²²

Külön probléma, hogy a MARC21 által a 005 rekordfej 18. a bibliográfiai leírás szabályai pozíción a lehetséges három érték (# *nem ISBD*, az AACR2 és *ISBD*) túlságosan szegényes. A HUNMARC itt hat típust enged meg (# *nem ISBD*, az AACR2, i *ISBD teljes*, p *ISBD egyszerűsített*, r *nem szabványos* és u *ismeretlen*). A 18. pozíció és a 40 \$e együttesen kell, hogy tájékoztassa a programozót arról, mire kell a konvertáláskor számítani, számítani kell-e, és hol mezőn belüli központoszási jelekre?

Amikor a MARC-ot mint adatforrást nem könyvtári háttérrel próbálja informatikus használni, hogy formálisan feldolgozza, óhatatlanul beleütközik ebbe a problémába. Ez különösen zavaró, ha az adatok felhasználása OPAC-on, vagyis az eredeti

kontextuson kívül történik (pl. a LOD felhőben, a Wikidatán, citációs formátumokban, az Euro-peanában stb.). Akár a MOKKA esetében is felmerülhet, hogy vajon leprogramozták-e, hogy a egyes anyagban az egyes rekordok megjelenítése figyelembe veszi-e a Tételfej 18. pozícióját, amelyben szigorúan meg kell(ene) adni az alkalmazott szabályzatot (és még mindig kérdés, hogy a rekord vajon tényleg követi-e az ott megadott módot).

A történetileg az „idők kezdetén” angol nyelvterületen kialakult központoszási gyakorlat valójában eleve idegen a MARC szellemiségétől, amelyben alapelv, hogy minden adatelem-típust elkülönítetten kezeljenek a mező–almező–indikátor szerkezet segítségével. Egykor elkezdett használata arra utal, hogy maguk a könyvtárosok nem értették igazán ezt a filozófiát. *Henriette Avram*nak a MARC korai történetéről írt művében olvasható²³, hogy a rendszer kezdeti célja a katalóguscédulák nyomtatása volt. Én teszem hozzá, hogy – ahogy ez lenni szokott – a praktikus cél elnyomta az elveket – és mire megjelentek a képernyőkön az online katalógusok, ez a gyakorlat beleégett sok könyvtárosi lélekbe.

A HUNMAC alkotói az 1990-es években eleve teljesen kiküszöbölték ezt a gyakorlatot. Egyébként újabban határozott törekvések vannak ennek az értelmetlen gyakorlatnak a felszámolására [3] és [13]. A MARC21 egyébként is folytonos fejlődésben van, az elmúlt évtizedben többek között számos új adatelem kapott benne helyet [14]. Az adattartalmának a gazdagságát illetően csodálatos rendszer. Ezt a gazdagságát valószínűleg a jövőben, amikor a gazdag adatokkal rendelkező rekordok jobban biztosítani fogják a pontosabb keresést, a mainál jobban fogják becsülni.

A másik probléma megoldása nagyobb feladat. A MARC-formátumnak nincs minőségbiztosítási rendszere. Ez utóbbi kérdéssel a továbbiakban nem foglalkozom, tárgyalása messzire vezetne; de ez sem megoldhatatlan probléma.

E néhány probléma miatt, sokszor eleve az értetlenség, a felkészületlenség, nagyon gyakran azonban a számítástechnikai autizmus hatására, a MARC létezésének jogosságát vonják kétségbe. Parádés példa erre *Roy Tennant* „MARC must die” című cikke [12]. A nagy szerzői bizonyosságra valló cím alatt – „a MARCnak pusztulnia kell” – azt írta, hogy „*The very nature of the MARC record is, to some degree, an anachronism.*” Azaz „a MARC-rekord bizonyos mértékig anakronizmus”. A „must

die” olvastán az ember elmélázhat azon, hogy honnan ez az indulat egy ilyen termékkel szemben...

Ez a sommás ítékezés mélységesen igaztalan. A MARC mező–almező–indikátor szerkezete egy klasszikus felszíni logikai szerkezet, amely a nem informatikus alkalmazók számára viszonylag könnyen és gyorsan megérthető. Ez alá a logikai szerkezet alá (vagy hozzá, ha tetszik) minden további nélkül kapcsolható egy olyan fordító vagy interfézszerkezet, amely átteszi a MARC szerkezetét a programozáshoz alkalmasabb szerkezetbe. Végül is a MARC XML-változata is ilyen. Ehhez képest a MARC említett felszíni logikai szerkezete sokkal inkább felhasználóbarát, ami a gyakorlati használóit illeti.

Tennant másik állítása, hogy „*Metadata situations like this are completely indefensible in the world of the web.*” Azaz „*az ilyen metaadat helyzetek az internet világában teljességgel védhetetlenek.*” Megint csak: azt mondja, hogy a MARC nem tud kezelni bizonyos (triviális és kevésbé triviális) mai felhasználási eseteket, és olyan metaadat-hordozóra lenne szükség, amely ezeket kezelni tudja. Hát ez az a bizonyos interfézszerkezet, mely a MARC említett felszíni logikai szerkezetét átfordítja. Se a MARC21-nek, se a feldolgozóknak ehhez semmi közük, szintisztán informatikai kérdés.

Ettől függetlenül van a MARC21-ben javítandó; a MARC tökéletesítésének útkeresése már húsz éve folyik. *Coyle* és *Thomale* közleményei is tanúskodnak erről. De akármilyen is lesz a jövőbeni MARC++, annak továbbra is a könyvtáros végfelhasználó által felfogható, az ő nem informatikus szakmájában értelmezhető logikai formátumban kell megjelennie. Más szóval egyszerűen kell tudni kiolvasni belőle az adatelemek függését, így például a jelenlegi MARC21-ben a mező-almező logikai szerkezetben megjelenített függéseket adatelem és részadatelem között. Ez az a logikai szint és formátum, amelyre a jövőben is szükségük van a feldolgozóknak. Se az XML, se hasonló formátumok erre nem alkalmasak.

Ez pedig megoldható.

Az igazán nagy problémákat az alkalmazás okozza. Ahogy Király Péter fogalmazott: „...*a legkiválóbb könyvtárak (Library of Congress, Harvard, Deutsche Nationalbibliothek etc.) sem tudnak hibátlanul dolgozni. Ez számomra döbbenetes.*”

A fenti összehasonlító elemzésben is kimutattam, hogy ennek a bajnak az oka nem a MARC szerkezete, hanem egyrészt a *gyakorlati alkalmazásának* hiányosságai. Másrészt, és ez a sokkal nagyobb baj: a *hibák többségét elsősorban a MARC-ot kezelő könyvtári feldolgozórendszerek (könyvtári szoftverek) súlyos elavultsága okozza.*

Irodalom

- [1] Avram, Henriette: MARC, its history and implications. Library of Congress, University of Michigan, 1975. 50 p.
<https://babel.hathitrust.org/cgi/pt?id=mdp.39015034388556&view=1up&seq=13>
- [2] Bilicsi Erika (szerk.): A MARC21 szerinti katalógizálás bevezetése az MTA könyvtár és információs központban. Budapest, 2018.
http://real-eod.mtak.hu/6918/1/marc21_VEGLEGES_pdf.pdf
- [3] Coyle, Karen: MARC21 as Data: A Start. = Code{4}lib Journal, 14. sz. 2011. 07. 25.
<https://journal.code4lib.org/articles/5468>
- [4] Draft PCC Guidelines for Minimally Punctuated MARC Bibliographic Records. Prepared by PCC Standing Committee on Applications. Approved by PCC Policy Committee, Sept. 2019.
www.loc.gov/aba/pcc/documents/PCC-Guidelines-Minimally-Punctuated-MARC-Data-draft.docx
- [5] HUNMARC. A bibliográfiai rekordok adatcsere formátuma. KSZ 4/1. 2002. március. 24
- [6] Király Péter: Metadata assessment for MARC records. Source code. 2017.
(<https://github.com/pkiraly/metadata-qa-marc>).
Some running instances: <http://134.76.163.21/dnb/>,
<http://134.76.163.21/cerl/>,
<http://134.76.163.21/gent/>.
- [7] Király Péter and Büchler, Marco. „Measuring completeness as metadata quality metric in European.” = 2018 IEEE International Conference on Big Data, IEEE, pp. 2711–2720.
<https://ieeexplore.ieee.org/document/8622487>. A modified version is available in the thesis as Chapter 2.
- [8] Király Péter: Measuring subject term usage in bibliographic records. 2019. 10. 30.
<http://pkiraly.github.io/2019/10/30/measuring-subject-term-usage-in-bibliographic-records/>
- [9] Király Péter. “Validating 126 million MARC records”. = DATeCH2019 Proceedings of the 3rd International Conference on Digital Access to Textual Cultural Heritage Brussels, Belgium – May 08–10, 2019. Published by ACM, 2019. ISBN: 978-1-4503-7194-0. pp. 161-168.

<https://dl.acm.org/doi/10.1145/3322905.3322929>. A modified version is available in the thesis as Chapter 4. (Slides: <http://bit.ly/qa-datech2019>)

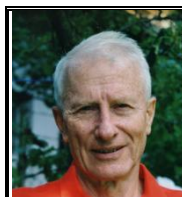
- [10] Király Péter: Record validation and quality assessment. Part II. 2019.
https://docs.google.com/presentation/d/e/2PACX-1vRU4J_rln00UVD7pNPT0_02NOad0HfSk_UKqRI_0v29y8QkMAplEDlyic0Ot_VE_paV6WBW29Fh_V-iN/pub?start=false&loop=false&delayms=3000#slide=id.g286e6f468f_0_66
- [11] MARC21 Format for Bibliographic data. Library of Congress. Update No. 30 (May 2020)
<https://www.loc.gov/marc/bibliographic/>
- [12] Tennant, Roy MARC must die. = Library Journal, 2010. 05. 21.
<https://www.libraryjournal.com/?detailStory=marc-must-die>
- [13] Thomale, Jason. Interpreting MARC: Where's the Bibliographic Data? Code4Lib Journal, 11. 2010. 09 21.
<https://journal.code4lib.org/articles/3832>
- [14] Ungváry Rudolf: A besorolási adatcsere-formátum bővülése. A legutóbbi két évtized fejleményei. = TMT. 2011. 09. 10., 58. évf. 9. sz., p. 371–386.
<https://tmt.omikk.bme.hu/tmt/article/view/868/10649>
- [15] Ungváry Rudolf: Besorolási, szabványosított, normatív vagy „autorizált”. = TMT, 2019. 06. 24., 66.évf., 6. sz., p. 328–342.
<https://tmt.omikk.bme.hu/tmt/article/view/12309/14064>
- [16] Ungváry Rudolf: MARC21 tartalmi adatmezők használata jelentősebb nagykönyvtárakban. Egy elemzés tanulságai. – Networkshop 2020. szeptember 2–4.
<https://nws.comp-rend.hu/program/> [online konferencia]

Hivatkozások

- ¹ A tanulmányt megelőzte egy előadás a NETWORKSHOP 2020. évi konferenciáján [16].
- ² Az MTAK 2019-es évi beszámolójában például szerepel:
https://konyvtar.mta.hu/download/dokumentumok/evi_jelentes/mtak2019.pdf
- ³ Szegedi Tudományegyetem,
<http://134.76.163.21/szte/>
- ⁴ A mesterséges nyelvű, szabályozott szótárakban származó egységek a kódszavak (a könyvtári gyakorlatban többnyire jelzetek), és a nyelvi rendszer, melynek szótárát alkotják, maga a kód.
- ⁵ A \$1 URI almezőbe a web teljes „világára” érvényes, tehát nem meghatározott rendszerben kiadott URI kerül. Ezért nevezik a MARC21-ben „real world object URI”-nak (azaz „valós hordozó URI”-jának). Az egyes rendszerekben is adhatnak ki URI-t besorolási rekordazonosító céljára, melyek a \$1 besorolási rekordazonosító vagy szabvány almezőbe kerülnek. Az angol megnevezés megtévesztő, mert az egyes rendszerekben kiadott URI-k is a valós világ dolgait azonosítják.
- ⁶ A táblázatok helyszűke miatt rövidítéseket kellett alkalmaznom. A MARC21 2018-as angol változatról készült egy „MARC 21: a bibliográfiai rekordok adatcsere-formátuma–MUNKAANYAG” című fordítás, mely csak az adatmezők, az almezők és indikátorok fordítását tartalmazza (ford. Kasza Zsófia.) <http://www.oszk.hu/sites/default/files/MARC21_bibliografiai_rekordok_adatcsere_formatuma.pdf>, és nem KSZ kiadás. A kéziratban nincs jelezve, hogy lektorálták volna, a fordításhoz egyáltalán nem vették figyelembe a már meglévő, 2002-es KSZ/4.1 fordítási megoldásait, a kéziratot a nyilvánosságban nem tárgyalták meg, a mező- és almezőnevek fordítása mereven követi az angol kifejezéseket. A benne alkalmazott magyar terminológiai problémáikról lásd [15].
- ⁷ Úgy tűnik, hogy ez nem közismert. A MOKKA specifikációjában például csak a 6XX-es mezőket tekintik tartalmi mezőknek. Holott a dokumentumtípusok elválaszthatatlanok a klasszikus, tárgyszavas és osztályozási jelzetek szerinti kereséstől. Önmagában dokumentumtípusra vajmi ritkán keresnek.
- ⁸ Az átszerkesztés fontosságát egyetlen magyar forrás említi [2] azzal összefüggésben, hogy a tipológiát belső használatra javították: „*Jobban strukturált, főleg a 007, 008 mezőkben figyelhető ez meg, illetve az új hordozók fizikai leírásának kibővült lehetőségeinél; a médiumok jobban besorolhatók a leginkább rájuk jellemző kategóriákba, pontosabban megadhatók a főbb jellemzőik.*” Ez a forrás részletesen tájékoztat a MARC21 mezőinek, almezőinek és indikátorainak többi változtatásáról, melyek túlnyomórészt az RDA-hoz való alkalmazkodás érdekében valósultak meg.
- ⁹ Túlnyomórészt a „Tartalmi jellemzők” a 24–17. pozíción.
- ¹⁰ Túlnyomórészt 9-es értéket tartalmazó lokális mezők.
- ¹¹ Hibás, sérült, nem előírt alakú ISBN/ISSN, ismételt nem ismételt, meghatározatlan (nem előírt értékű).
- ¹² HUNMARC=ár
- ¹³ Tudomásom szerint 2019-ben volt is egy nagy ellenőrző szűrés külső szoftverrel (mivel a könyvtári szoftver erre nem ad lehetőséget), és a kimutatott hibák jelentős részét heroikus munkával az OSZK szakemberei kigyomlázták. Ennek tanulságai is hozzájárulnának a szakma magyar színvonalának eme-

- léséhez – és az eddig könyvtári szoftverekkel való leszámolásnak.
- ¹⁴ A gondosabb feltárás lehet az oka. Majdnem 100 %
- ¹⁵ Volt és van talán ma is olyan magyar könyvtári felsővezető, aki a legszívesebben megszüntetné az ETO használatát. Még rosszabb az az eset, ha valaki az ETO „egyszerűbb”, azaz lebutított változatának alkalmazását akarja, ami az ETO teljes meg nem értését jelenti. Remélhetőleg a könyvtári szoftverek újabb generációinak megjelenésével, amely könnyebbé teszi majd az ETO kereséskori használatát, az ilyen szakmaiatlan szándékok elenyésznek.
- ¹⁶ Nem azonos a 655 formai alosztással.
- ¹⁷ HUNMARCban „*rokonsági kiegészítő*”.
- ¹⁸ Ráadásul az egyes magyarországi települések, lakott helyek, geomorfológiai képződmények (hegyek, dombok, patakok stb.) geokoordinátáit a KSH nyilvántartja, és különösebb nehézségek nélkül letölthetők lennének azoknak a helyeknek a geokoordinátái, melyek az adott könyvtár online adatbázisában szerepelnek. Ily módon később, amikor az online katalógusból adott helynévről át lehet majd térni a hely térképi megjelenítésére, ezek már eleve rendelkezésre állnának.
- ¹⁹ Hogy kerülhet az ellenőrzött MOKKA-ba egy ennyire abszurd mezőhívójel?
- ²⁰ Szakmai szlengben ezt a fajta tárgyszót szabad tárgyszónak nevezik, ami nem utal pontosan arra, hogyan, milyen feltételekkel lehető fölvenni.
- ²¹ Hogy lehet szabadon (tehát nem kötött szótárból, tárgyszójegyzékből stb.) fölvett tárgyszó esetében besorolási rekordazonosító és forrás?
- ²² A ma végre az angol szakirodalomban észrevett, almezőn belül feltüntetett központozási jelek helytelen gyakorlatáról már a múlt század végén (tehát kb. 30 éve) a magyar szakirodalomban egy sor, a gyakorlat abszurditását tárgyaló cikk jelent meg.
- ²³ A teljes szöveg hozzáférhető [1]. Érdemes elolvasni, és megnézni a bibliográfiáját is: mi minden játszódott le már 1966–1967 környékén a Kongresszusi Könyvtár világában. Ennek fényében értékelhető a teljesítmény, hogy az OSZK-ban Sipos Márta és Szűcs Jenőné szervező és rendszertervező munkája nyomán, a Számítástechnikai Koordinációs Intézet néhány munkatársának közreműködésével 1978-ra el tudott indulni a MNB számítógépes rendszere és kiadvány-előállítása. (Szűcs Jenőné: A nemzeti bibliográfia számítógépes előállítása és az erre építhető új szolgáltatások. In: A Magyar Könyvtárosok Egyesülete évkönyve 1978. Bp. 1979. 25. p.)

Beérkezett: 2020. X. 23-án.



Ungváry Rudolf
informatikus mérnök.
E-mail: ungvaryr@gmail.com