

Rendező:



Networkshop 2020
ONLINE

Házigazda:



PÉCSI TUDOMÁNYEGYETEM
UNIVERSITY OF PÉCS

„Esélyeink és kihívásaink a digitális transzformáció világában”

Országos Online Konferencia
2020. szeptember 2–4.



INNOVÁCIÓS ÉS TECHNOLÓGIAI
MINISZTERIUM



Szerkesztette: Tick József, Kokas Károly, Holl András

Tipográfia és tördelés: Vas Viktória

Networkshop
2020. szeptember 2-4. Pécsi Tudományegyetem, (On-line)
konferencia előadásainak közleményei

ISBN 978-615-01-0376-1
DOI: [10.31915/NWS.2020](https://doi.org/10.31915/NWS.2020)

Kiadja a HUNGARNET Egyesület
az MTA Könyvtár és Információs Központ közreműködésével
Budapest
2020

Borítókép: [freepik.com](https://www.freepik.com)

MARC21 tartalmi adatmezők használata jelentősebb nagykönyvtárakban Egy elemzés néhány tanulsága

Ungváry Rudolf
Országos Széchényi Könyvtár
ungvaryr@gmail.com

1. Forrás és tárgy

Király Péter letöltötte és elemezte 16 jelentős nemzeti és tudományos könyvtár (köztük két magyar nagykönyvtár és a MOKKA) bibliográfiai MARC21 állományát [4]–[10].¹ Az OSZK katalógusállományának letöltéséhez nem járultak hozzá.

Az elemzés kiterjedt a MARC21 (és adott esetben HUNMARC) összes adatelem-típusára és -értékére. A 1. ábrán² a 650-es mező első néhány almezőjének elemei (almezőjel, név stb.) láthatók.

path	label	terms	records	
			count	%
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. Megjeleníthetők a bibliográfiai rekordok, az almező tartalma, lekérhető akár a teljes, akár a MARC21 formátuma, Látható az adott almezőben rögzített rekordok száma, és ezek %-a az összes rekordhoz képest.

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

1 Lásd a <http://134.76.163.21/.../>, ahol a /.../ helyébe a források rövidítése kerül (pl. SZTE)

2 Szegedi Tudományegyetem, <http://134.76.163.21/szte/>



Az elemzés alapján következtetések vonhatók le arról, hogy milyen főbb feldolgozási jellegzetességei vannak az egyes könyvtárak katalógusainak, az alkalmazott könyvtári feldolgozórendszeréről és kötött szótárakról, valamint összehasonlíthatók az egyes könyvtárak katalógusai egymás között.

A tanulmányban csak a források *tartalomra vonatkozó* adatait elemeztük. Csak HUNMARC-almezőket halványan, a hibás almezőket halványan és három kérdőjellel jelöltem.

2. A hibák összehasonlítása

Király felmérésében az „Issues” fül alatt („Hibatáblázatok”) hibatípusok szerint ellenőrizhetők a konkrét hibák. Túlnyomó többségük megfelelő hibaellenőrző szoftver esetén elkerülhető lenne.

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										
hibás érték ³	11238	3889	864505	333635	281454	110011	3656937	363434	20317130	5240827
egyéb vezérlő mezők és almezők										
nem ismételt mező ism.	740	740	65	67	8	8	240	240	4705	4698
nem előírt hívójelű mező ⁴	-	-	-	-	308347	268696	2884	2626	1115714	479623
\$ind hibás indikátor	53161	8066	3031	5066	192690	178461	5015	4271	3541444	2298715
\$X hibás almező ⁵	2368900	2483725	430833	529103	87	57	31197	26900	34759334	8864772
nem előírt hívójelű alm.	2.478770	2365446	429314	520863	520863	429314	11417	8048	334650380	8861086
ebből 65X almezői	728	768	16	16	21	21	75	71	682	624
nem ism. alm. ismétlődése	-	-	-	-	6452	3867	1900	1857	881787	776565

1. táblázat. Pozíció-, mező-és almezőhibák.

A könyvtárak egy részében (így az SZTE-ben) a 008-as mezőben nem töltnek ki minden adatalemet képviselő pozíciót, és vannak pozíciók, melyek tartalmát automatikusan generálják. Sok hiba származhat az automatizált betöltésekből.

Az 5. sorban a „nem ismételt 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 következetlen (zajos) képet

3 Túlnyomórészt a „Tartalmi jellemzők” a 24–17. pozíción.

4 Túlnyomórészt 9-es értéket tartalmazó lokális mezők.

5 Hibás, sérült, nem előírt alakú ISBN/ISSN, ismételt nem ismételt, meghatározatlan (nem előírt értékű).

mutatnak, hogy meglehetősen kis valószínűséggel lehetnek tervezettek. A 65X mezők hibái pedig elsősorban beviteli hibák..

A MOKKA közös katalógusban feltűnő, hogy a szűrések ellenére mennyi hiba került át. A szűrőprogramok, valamint a lokális mezők dokumentálása közismerten általában nem valami tökéletes, és ami van, nehezen hozzáférhető, és ez a MOKKA esetében is így lehet. Egyedül az OSZK által átadott állomány szűrési specifikációja ismert.⁶

Mutatóban néhány hiba az SZTE állományából, melyekhez hasonlók más könyvtárakban is előfordulnak.

<u>hibatípus:</u>	<u>hiba:</u>	<u>rekord db:</u>	<u>előfordulás db:</u>
Hibás 008-as mező /720 változat)		3656937	363434
Hibás jelzet (158 változat):		4807	4556
Pl. 008/18-21	,r' in ,ar'	3392	3392
<i>Azaz a könyvek 18–21. pozícióin az egy karakteres ,a' (illusztráció) érték helyett hibásan ,ar'szerepel.</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 kellene lennie..</i>			
Hibás almező		31197	26900
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.</i>			
Almező: hibás ISBN (10348 változat)		13115	13012
020\$a	A ,0238-3349(fűzött)' nem ISBN-formátum	8	8

E mezők hibái megelőzhetőek lennének, ha a könyvtári rendszerek lennének jobban felkészítve ezek ellenőrzésére, hibaüzeneteire.

6 Az OSZK hibáinak kiszűrése: „Szabó Julianna: MOKKA inicializálás. Az OSZK rekordok szintaktikai ellenőrzése. Egykötetes és többkötetes közös adatokat leíró rekordok. 2003. 02. 07. – 92 p. [kézirat]”.



3. A OXX számozások és jelzetek mezőcsoportja

A kitöltöttségük képet ad az egyes könyvtárakban folyó feldolgozás teljességéről.

3.1 A bibliográfiai rekordazonosítók

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
010\$a LC azonosító	5894	0.03	256940	14.56	1	0.00	161	0.01	56366	0.63
015 nemzeti bibliográfiai azon-										
\$2 forrás	11345731	59.63	26219	1.18	-	-	4	0.00	593	0.01
\$a azonosító	11345731	59.63	71517	1.49	-	-	7	0.00	6284	0.07
\$z törölt azonosító	1995499	10.49	637	0.04	-	-	-	-	9	0.00
016 bibl. intézmény azon.										
\$2 forrás	19 mill.	100	31802	1.80	-	-	5	0.00	840	0.01
\$a azonosító	19 mill.	100	32155	1.82	-	-	6	0.00	844	0.01
019\$a [helyi adatelemtípus]	-	-	2	0.00	-	-	-	-	-	-

2. táblázat. A OXX mezőcsoport bibliográfiai rekordazonosítói.

Egy nemzeti könyvtárba (DNB) alig kerül más nagykönyvtárból származó rekord, mely a 010\$a LC-azonosítót tartalmazza, a DNB 0.03%-nyi LC-azonosítót tartalmazó rekordja kivétel. A GENT-ben viszont láthatóan több ilyen rekord is van (14.56%).

Az adott könyvtár 016 *nemzeti bibliográfiai intézmény azonosítójának* a kitöltöttsége a DNB esetében érthető, hogy teljes. Ez abból is látható, hogy értékeinek száma (19 millió) megegyezik a 4. táblázatban látható 040 *katalogizálás forrása* mező értékeinek a számával – ami összehasonlításban nagyon gondos feldolgozásra vall. Ez vonatkozik 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 nemzeti bibliográfiai azonosítóval. A SZTE-ben ezeket a mezőket nem használják, és lehet, hogy az MTAK-ban sem. A letöltés során bekerült értékek automatizáltan könnyen törölhetők.

A DNB esetében csak a katalógusrekordok 59.63 %-a kerül be a nemzeti bibliográfiába. (Az OSZ esetében kb. 65 % könyveknél az átlagos bekerülési arány.)

3.2 A kiadvány-azonosítók

A két kiadvány-azonosító az ismeretlen almezőjelek és ezek almező-tartalmait illetően afféle állatorvosi ló. A hibák abszolút száma ugyan elhanyagolható, nem vonható le belőlük az érintett katalógusok minőségéről általános következtetés, de lehangoló bizonyítványt ad a katalogizáló szoftver ellenőrzési képességeiről.

Ilyen zaj a külföldi könyvtárakban meglehetősen ritka. A jobb minőségű szoftver hiánya meg anyagi. A „nemzet csinosítása” alulfinanszírozott.

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 nezőkapcs. és soroz.	-	-	3	0.00	-	-	6	0.00	35	0.00
\$9 [helyi almező]	5759980	30.27	-	-	-	-	91	0.01	22	0.00
\$C ???	-	-	-	-	-	-	-	-	3	0.00
\$a ISBN azonosító	5721373	30.07	621661	35.24	217607	20.68	645528	50.86	2738551	30.81
\$b ???	-	-	-	-	2	0.00	23	0.00	457	0.00
\$c hozzáférés feltétele ⁷	7044386	37.02	9950	0.56	45	0.00	25153	1.98	30041	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	3880	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	-	-	87028	4.93	8095	0.77	6	0.00	4524	0.05
\$r ???	-	-	-	-	-	-	-	-	2	0.00
\$s ???	-	-	-	-	-	-	3	0.00	4	0.00
\$t ???	-	-	-	-	-	-	2	0.00	1	0.00
\$z törölt ISBN	85751	0.45	8869	0.50		1187	4283	0.34	47795	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	-	-

⁷ 1974 előtt a MARC21-ben \$c = ár volt, a HUNMARC-ban továbbra is ár.



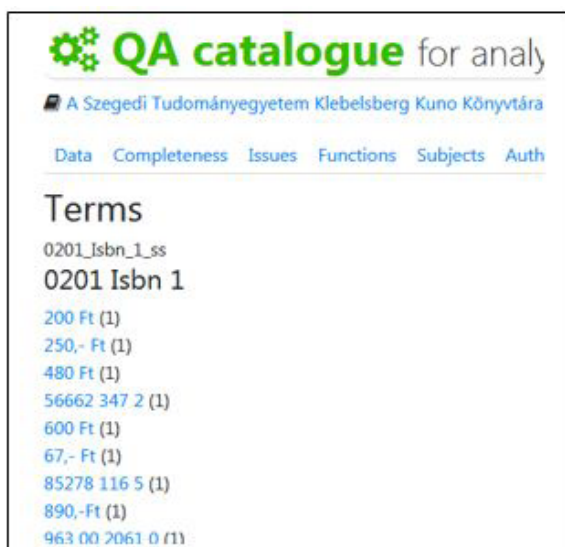
\$6 kapcsolódás	-	-	-	-	1	0.00	-	-	1	0.00
\$7 ???	-	-	-	-	1	0.00	-	-	-	-
\$8 mezőkapcs. és sorsz.	-	-	-	-	-1	0.00	-	-	-	-
\$S ???	-	-	-	-	1	0.00	-	-	-	-
\$a ISSN azonosító	1.651 mill.	8.68	427	0.02	19357	1.84	69743	5.49	85055	0.96
\$b ???	-	-	-	-	-	-	-	-	1	0.00
\$c ár	1343	0.01					28	0.00	872	0.01
\$f ???	-	-		-	-	-	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	40582	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

3. táblázat. Az ISBN és az ISSN.

A halványan írt ismeretlen almezők vagy értelmezhetetlen, vagy sajátos ISBN ill ISSN értékeket, vagy az árat tartalmazzák, miközben csak a HUNMARC szerint mind az ISBN-ben, mind az ISSN-ben létezik az önálló \$c ár almező is. A sok ismeretlen almezőnek bizonyára lehet magyarázata. A példaként választott külföldi könyvtárak esetében, ilyen zajnak tűnő jelenséget nem találni.

Önmagában is különös, hogy lehetséges a formátumokban nem létező almezőket egyáltalán rögzíteni és ezeknek értéket adni anélkül, hogy a rendszer legalább üzenetet ne küldene erről a tényről. Egy biztos: a könyvtári feldolgozórendszer megfelelő hibaellenőrzési 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 állományában.

3.3 Néhány további speciális azonosító

A speciális azonosítókat tartalmazó almezők közvetve a tartalom szerinti válogatást tennék lehetővé. Ez messze nincsen igazán felhasználóbarát módon biztosítva a katalogizáló rendszerekben.

MARC21 mező/ adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
024\$a egyéb szabv. azon.	8.479 mill.	44.56	103619	5.87	1644	0.16	867	0.07	97762	1.10
028\$a zenemű kiadói azon.	1408	7.40	7391	0.42	-	-	-	-	1292	0.01
030\$a CODEN	6,386	0.03	37	0.00	-	-	12108	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	283252	16.06	774923	73,64	864342	68.10	7062486	79.45
040 katalogizálás forrása										
\$a eredeti intézmény	19 mill.	100	1764209	100	1052265	100	1268276	100	7062997	79.45
\$b katalogizálás nyelve	19 mill.	100	-	-	1052286	99.99	972901	76.65	541	0.01
\$c átíró intézmény	19 mill.	100	-	-	1052202	99.99	4.470	0.35	176	0.00
\$d módosító intézm.	19 mill	100	-	-	119379	11.34	237527	18.71	277885	2.56
\$e leírási megállapodások	606123	3.19	-	-	22	0.00	10	0.00	4	0.00

4. táblázat. A speciális azonosítók és a katalogizálás forrása.



Figyelemre méltó, hogy egyedül a DNB-ben szerepelnek – a teljes állományhoz képest elhanyagolható mennyiségben – geokoordináta értékek, holott ezek idővel, amikor a katalógustételekben szereplő helynek magából a katalógusból kilépve majd a weben térképi helyként megjeleníthetők lesznek, alapvetőek.

A SZTE 035\$a rendszerazonosítói OCLC-azonosítók.

A 040-es mezőt minden rekordnak tartalmaznia kell. A \$d módosító intézmény és a \$e leírási megállapodások almező kivételével a 040-es mező többi \$a, \$b, \$c almezőjének ugyanannyi értéket kellene tartalmaznia. 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 keletkezik, az betöltés következménye. A GENT esetében, talán hogy elkerüljék az összhangra törekvés nehézségeit, eleve csak a 040\$a eredeti intézmény adatát adják meg.

A viszonylag gyengén kitöltött 040\$e leírási megállapodások tartalma 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 például használnak-e központosítási jeleket valamelyik – például a 300 főcím adatmezőben?

3.4 A nyelv

Ebben a mezőben a magyar könyvtárakban ugyancsak sok az ismeretlen almező, és értékeik többnyire a nyelvek jelzetei („kódjai”). Lehet, hogy a 041 nyelv mező esetében is helyileg használt almezőkről van szó, lehet, hogy tévesen lettek rögzítve. Vagy itt is lehetséges, hogy a letöltött rekordok okozták a „fertőzést”?

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	-	-	67268	0.76
\$7 ???	-	-	-	-	1	0.00	-	-	-	-
\$8 mezőkapcsoló és sorsz.	42000	0.22			-	-	-	-	-	-
\$a nyelv	15.7 mill.	82.58	101564	5.76	993704	94.43	148840	11.73	4126631	46.42
\$b összefoglalás nyelve	-	-	4791	0.27	14261	1.36	15930	1.26	67268	0.76
\$c ???	-	-	-	-	-	-	6	0.00	18	0.00
\$d ének vagy szöveg ny.	-	-	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

MARC21 mező/ adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
\$f tartalomjegyzék nyelve	-	-	100	0.01	1095	0.10	1680	0.13	10677	0.12
\$g egyéb járulékok nyelve	-	-	119	0.01	2439	0.23	3212	0.25	14527	0.16
\$h eredeti nyelve	569679	2.99	12014	0.68	55266	5.25	70562	5.56	367972	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 járulék 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
\$q ???	-	-	-	-	-	-	-	-	1	0.00

5. táblázat. A nyelvi azonosítók

Egyetlen katalógusban sincs minden rekord esetében megadva a nyelv. 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 DNB-nek megfelelő színvonalú.⁸ 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 helyei katalogizálási szabályzatok különböző értelmezései magyarázhatják, hogy hol milyen további nyelvi almezőket használnak. A hibás almezők automatizáltan könnyen törölhető.

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 meg a gyakorlatban, és ha nem, miért.

⁸ Tudomásom szerint 2019-ben volt is egy nagy ellenőrző szűrés külső szoftverrel, és a kimutatott hibák jelentős részét heroikus munkával az OSZK szakemberei kigyomlálták. Ennek tanulságai is hozzájárulnának a szakma magyar színvonalának emeléséhez.



3.5 A OXX mezőcsoport tartalmi mezői a 08X mezők kivételével Bennük az értékek kötött szótárból származó jelzetek vagy tárgyszavak.

MARC21 mező/ adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
043 földrajzi területazonosító										
\$2 helyi területazon. forrása	-	-	2	0.00	-	-	-	-	-	-
\$a földrajzi területazonosító	-	-	91420	5.18	-	-	14	0.00	8126	0.09
\$b helyi területazonosító	-	-	28	0.00	-	-	-	-	2	0.00
\$c ISO kód.	536976	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 országcód	-	-	-	-	3	0.00	18	0.00	733150	8.25
\$b részterület helyi azonosítója	-	-	556	0.03	-	-	1	0.00	1	0.00
\$c ISO országcód	18.8 mill. ⁹	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	17625	1.67	-	-	9	0.00
072 átfogó szakterület (kategória) jelzete.										
\$2 forrás	-	-	5581	0.32	17423	0.66	-	-	27	0.00
\$a átfogó szakterület	-	-	5949	0.34	17625	1.67	3	0.00	3423	0.04
\$x átfogó részterület	-	-	453	0.03	-	-	-	-	-	-
090										
\$6 ???	-	-	-	-	-	-	1	0.00	-	-
\$7 ???	-	-	-	-	-	-	1	0.00	-	-
\$a ???	-	-	-	-	-	-	11370	16	-	-
\$b ???	-	-	-	-	-	-	16	0.00	-	-
\$i ???	-	-	-	-	-	-	4	0.00	-	-
\$a ???	-	-	-	-	-	-	1	0.00	-	-
\$b ???	-	-	-	-	-	-	1	0.00	-	-

6. táblázat. A OXX mezőcsoport tartalmi azonosítói a 08X mezők kivételével.

⁹ A gondosabb feltárás lehet az oka. Majdnem 100%

Ezeket a tartalomra vonatkozó jelzeteket, ill. 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.

A MOKKÁ-ban feltűnő a sok az – egyébként nem kötelező – 044\$a MARC országjelzet, a \$c ISO országjelzet viszont szinte üres. Nyilván nem volt utánkötetés. 2001-től ugyanis a MARC21 bevezette a 044\$c almezőt az ISO-országjelzetek számára, és ettől kezdve a \$a almezőt a MARC országjelzetek számára van fenntartva. A DNB-ben az utánkötetés megtörtént (és az OSZK-ban is).

3.5 A 08X tartalmi mezők

A 080, 082, 083, 084 és 085 mezők értékei klasszikus osztályozási rendszerek jelzetei. 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.

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 ¹⁰ a	-	-	-	-	1	0.00	28	0.00	89	0.00
\$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	-	-	100245	5.68	535691	50,91	1038987	81.85	3566807	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

10 A \$0 almezőbe a web teljes „világára” érvényes, tehát nem meghatározott rendszerben kiadott URI kerül. Az egyes rendszerekben is adhatnak ki URI-t besorolási rekordazonosító céljára, melyek a \$1 almezőbe kerülnek.



MARC21 mező/ adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
\$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.	7365077	38.70	112572	6.38	-	-	50	0.00	37461	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	7365118	38.71	192905	10.39	1	0.00	745	0.06	82915	0.93
\$b Cutter példányazon.	-	-	4082	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	-	-	-	-	-	-
083a további TO jelzet	-	-	582	0.03	-	-	7	0.00	21	0.00
084\$a egyéb oszt. jelzet	4602135	24.18	52588	2.98	1	0.00	5136	0.40	87	0.00
085\$a szintetizált oszt. jelzet	1348402	7.09	-	-	-	-	3	0.00	8	0.00

7. táblázat. A OXX mezőcsoport tartalmi azonosítói.

A DNB-ben a rekordjainak jelentős része tartalmazza a TO-jelzetet. A DNB gyakorlata reálisan csak az OSZK gyakorlatával volna összehasonlítható. Egészében úgy tűnik, hogy a külföldi könyvtárakban valamivel több osztályozási rendszert alkalmaznak, mint Magyarországon.

Minden jel szerint 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. Technikailag semmiféle nehézséget nem jelent a megvalósítás.¹¹

Amagyar könyvtárakban az ETO alkalmazásának kultúrája közel évszázados hagyományra tekinthet vissza, és ezzel magyarázható, hogy a használata mindegyik könyvtárban jelentős.

A hibás ETO-almezők és a nem használt TO letöltések következtében keletkezett almezők javítása időközben a SZTE-be megtörtént. A könyvtári szoftver jobb ellenőrzési adottságai esetén ilyen hibák eleve elkerülhetők lennének.

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 pedig 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.

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 nevéken 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.

MARC21 mező/ adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
600 személynév tárgyszó										
\$# ???	-	-	-	-	-	-	-	-	193	0.00
\$0 besorolási rekordaz.	378007	1.99	3860	0.22	3	0.00	1	0.00	12	0.00
\$1 URI	-	-	1	0.00	-	-	-	-	6	0.00
\$2 forrás	433870	2.27	79207	4.49	1	0.00	63060	4.97	73665	0.83

11 Volt és van talán van ma is olyan magyar könyvtári felsővezető, aki a legszívesebben megszüntetné, az ETO használatát. Még rosszabb az ETO „egyszerűbb”, azaz lebutított változatának a szándéka. Ez az ETO teljes meg nem értése. A könyvtári szoftverek újabb generációinak megjelenésével, amely könnyebbé teszi majd az ETO kereséskori használatát, az ilyen szakmaiatlan szándékok elenyésznek.



MARC21 mező/ adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
\$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.	18733	0.10	-	-	1	0.00	-	-	4	0.00
\$D ???	-	-	-	-	-	-	-	-	1	0.00
\$M ???	-	-	-	-	-	-	-	-	1	0.00
\$a rendszó (személynév)	433676	2.28	79207	4.49	172001	16.35	64228	5.06	490403	5.52
\$b uralk. sorszám	88640	0.05	4343	0.25	5582	0.53	3788	0.30	27158	0.31
\$c méltóság, foglalkozás	39799	0.21	16501	0.94	29909	2.84	8245	0.65	125575	1.41
\$d kronologikus kiegészítő	351243	1.85	61884	3.51	139159	13.22	6592	0.52	391773	4.41
\$e kapcsolatjelölő	-	-	67	0.00	23725	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	-	-	2313	0.03
\$k formai alcím ¹²	-	-	36	0.00	-	-	-	-	3	0.00
\$l mű nyelve	-	-	61	0.00	1	0.00	-	-	28	0.00
\$m zenei változat hordozója ¹³	212	0.00	15	0.00	238	0.02	-	-	416	0.00
\$n mű részének száma	819	0.000	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	-	-	2297	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	36496	0.19	9398	0.53	10790	1.03	50	0.00	13451	0.15
\$' ???	-	-	-	-	-	-	-	-	1	0.00

8. táblázat. A 600-as mező összehasonlítása.

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

¹² Nem azonos a 655 formai alosztással.

¹³ HUNMARCban „rokonsági kiegészítő”.

(Gemeinsame Normdatei), miáltal a 600-as mezőbe kerülő nevek többségének ez lett a forrása. Hasonló lehet a helyzet a GENT-ben is. A SZTE-ben ugyancsak magát a könyvtárat adják meg forrásként.

A szintén új, az egész webre kiadott \$1 URI almezőt (*real world object URI*) pedig még alig használják. Mindkét almező új, 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.

Különös jelenség, hogy a DNB \$2 forrás almezőjében számszerűen egész kicsit több a forrásértékek száma, mint a \$a rendsző almezőben a rendsző-értékeké. Megint csak valószínűsíthető, hogy ebben elsősorban a szoftver a ludas, amely ezt jelzés nélkül megengedi.

Az MTAK katalógusában használják viszonylag a leginkább a személynév tárgyszót (\$a rendsző 16.35%). Ennek a könyvtárnak néhány más almezőjében is (\$c uralkodói sorszám és \$d méltóság) az átlagosnál több az érték. A SZTE-ben a \$2 forrás almező kitöltöttsége majdnem összhangban van egymással (a teljes azonosság érdekében a javítás folyamatban van). A magyar könyvtárakra a leginkább jellemző a \$c és \$d használata.

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

4.2 A 648-as kronologikus tárgyszó

A kronologikus tárgyszavak elkülönített minősítése szorosan összefügg a 045 rekord tartalmának időtartománya és 046 különlegesen kódolt dátumok mezőkkel, és azok összhangolt használatával.

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 rendsző (kronologikus)	439379	2.31	4976	0.28	-	-	-	-	519	0.01

9. táblázat. A 648-as mező összehasonlítása.

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 SZTE katalógusában – és feltételezhetően az MTAK katalógusában is – a 650-es mező \$y kronologikus altárgyszavát használják helyette. A MOKKÁ-ban megjelenő értékek történelmi szakkönyvtárból származnak. Mivel a Köztauruszban a kronologikus tárgyszavak eleve minősítve vannak a 148-as mezőben, az OSZK katalógusában kényszerűen a 648-as mezőbe kerülnek.



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, ugyan úgy mint a 046 esemény, tartalom dátuma mezők teljesen hiányzanak. Pedig ezek a tartalmi keresés szempontjából nagyon fontosak lennének.

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.

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 feltárásról.

MARC21 mező/ adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
650 szaktárgyszó										
\$# ???	-	-	-	-	-	-	-	-	7118	0.08
\$\$???	-	-	-	-	-	-	-	-	74	0.00
\$0 besorolási rekordazon.	2844292	14.95	67743	3.84	-	-	-	-	72	0.00
\$1 URI	-	-	-	-	-	-	-	-	53	0.00
\$2 forrás	3612054	18.98	689532	39.08	64	0.01	733857	57.82	1402485	15.78
\$3 dokumentumrész	-	-	-	-	-	-	1922	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.	209766	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)	3612054	18.98	689532	39.08	140030	13.31	733809	57.82	3119811	35.10

MARC21 mező/ adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
\$b földr. nevet köv. tárgyszó	-	-	50	0.00	12	0.00	26	0.00	13808	0.16
\$c esemény helye	-	-	2690	0.15	13	0.00	17	0.00	1716	0.02
\$d esemény időtartama	-	-	3363	0.19	126	0.00	35	0.00	3261	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	208211	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
\$' ???	-	-	-	-	-	-	-	-	1	0.00
\$v formai altárgyszó	-	-	63791	3.62	43158	4.10	43158	4.10	64029	0.72
\$x általános altárgyszó	334101	1.76	312975	17.74	120712	11.47	120712	11.47	1447133	16.28
\$y kronologikus altárgyszó	-	-	61933	3.51	70134	6.66	70134	6.66	501504	5.64
\$z földrajzi altárgyszó	-	-	173050	9.81	65392	6.21	65392	6.21	369933	4.16

10. táblázat. A 650-es mező összehasonlítása.

A DNB-ben a \$0 besorolási rekordazonosítónak gyakrabban van értéke (14.95%), mint a 600-as mező esetén, és ez a helyzet a \$a rendszó esetén is, melynek értékei teljesen összhangban vannak a \$2 forrás almező értékeivel. Ez bizonyára a besorolási tárgyszó-rendszerének (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 rendszó 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.

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 feltárása utal. Ez figyelhető meg a Finn Nemzeti Könyvtár, és tapasztalatom szerint ez az OSZK esetében is.

A sok értelmezhetetlen – valószínűsíthetően hibás – almező kiábrándító. Nem a figyelmetlenség miatt, mert a hibák mennyisége hibahatáron belüli, hanem a könyvtári szoftver ellenőrző képességeinek tükrében.

4.4 A 651-as földrajzi tárgyszó

A földrajzi tárgyszó a keresés szempontjából gyakran szervesen kiegészíti a szaktárgyszavak által megnevezett fogalmakat.

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
651 tárgyszó										
\$0 besorolási rekordazon.	984088	5.17	12199	0.69	-	-	-	-	12	0.00
\$1 URI	-	-0	1	0.00	-	-	-	-	8	0.00
\$2 forrás	1114929	5.86	84056	4.76	1	0.00	35122	2.77	66400	0.75
\$3 dokumentumrész	-	-	-	-	-	-	113	0.01	1542	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.	41490	0.22	-	-	-	-	-	-	6	0.00
\$9 ???	728	0.00	-	-	-	-	-	-	-	-
\$@ ???	-	-	-	-	-	-	-	-	1	0.00
\$M ???	-	-	-	-	-	-	-	-	2	0.00
\$a rendszó (tárgyszó)	1114929	5.86	84056	4.76	7325	0.70	50893	4.01	674195	7.58
\$b tárgyszót követ. földr név	-	-	4	0.00	4	0.00	8	0.00	3495	0.04
\$c kiegészítő elem	-	-	-	-	6	0.00	2	0.00	10042	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	78322	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

11. táblázat. A 650-es mező összehasonlítása.

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 nincsen összhang. (Az OSZK-ra vonatkozóan Köztaurusz használata következtében a 648-as, 650-es és 651-es mezőkre 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 használ kötött szótár igénybevétele nélkül, ezektől teljesen függetlenül vehető föl (nevezik „index kifejezésnek is, de mutatószó bármelyik 6XX-es mező tárgyszava lehet). Más szóval eleve nem is lehet olyan tárgyszó, mely egy ilyen kötött szótárból származnék.¹⁴

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

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
653 szabadon választott tárgyszó										
\$0 besorolási rekordaz ??? ¹⁵	-	-	-	-	-	-	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 rendszó (szab. v. tárgyszó)	6116045	32.14	-	-	68550	6.51	122	0.01	327527	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

12. táblázat. A 653-as mező összehasonlítása.

14 Szakmai szlengben ezt a fajta tárgyszót szabad tárgyszónak nevezik, ami nem utal pontosan arra, hogyan, milyen feltételekkel lehető főlvenni.

15 Hogy lehet szabadon (tehát nem kötött szótárból, tárgyszójegyzékből stb.) főlvett tárgyszó esetében besorolási rekordazonosító és forrás?



Az összes többi mező esetében hibákról van szó – és a szoftver hiányos ellenőrzési képességeiről.

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.

MARC21 mező/adatelem	DNB	%	GENT	%	MTAK	%	SZTE	%	MOKKA	%
655 formai tárgyszó										
\$0 besorolási rekordazon.	5087226	16.74	15908	0.85	-	-	-	-	6	0.00
\$2 forrás	5098459	26.80	197768	11.21	62	0.01	1493	0.12	260013	2.92
\$3 dokumentumrész	-	-	-	-	-	-	-	-	1	0.00
\$4 ???	-	-	-	-	1	0.00	-	-	1	0.00
\$a rendszó (formai tárgyszó)	5098603	26.80	197768	11.21	19226	1.83	1493	0.12	465447	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

13. táblázat. A 655-ös mező összehasonlítása.

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 650 mező \$v formai altárgyszavát használják helyette.

5. Kitekintés a MARC21-re, és a tapasztalatok általánosítása

A MARC21 változóban van, egy jövőbeli MAARC++-ra továbbra is szükség lesz [12] [2]. A számítástechnikai szempontokra korlátozódó szemlélet, afféle „szakmai autizmus” következtében születnek a MARC21 megsemmisítését javasoló, méltánytalan közlemények is [11]. Minderre, és az összehasonlító elemzés további részletkérdéseire, a tanulások általánosítására egy másik közleményben térek ki. [15].

6. Bibliográfia

- [1] Bilicsi Erika (szerk.): [A MARC21 szerinti katalogizálás bevezetése az MTA könyvtár és információs központban. Budapest, 2018. http://real-eod.mtak.hu/6918/1/marc21_VEGLEGES_pdf.pdf](http://real-eod.mtak.hu/6918/1/marc21_VEGLEGES_pdf.pdf)
- [2] Coyle, Karen: MARC21 as Data: A Start. In: Code4lib Journal, 14. sz. 2011. 07. 25. <https://journal.code4lib.org/articles/5468>

- [3] 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
- [4] HUNMARC. A bibliográfiai rekordok adatcsere formátuma. KSZ 4/1. 2002. március.
- [5] 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/>.
- [6] Király Péter and Büchler, Marco. “Measuring completeness as metadata quality metric in Europeana.” In: 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.
- [7] 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/>
- [8] 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>)
- [9] Király Péter: Record validation and quality assessment. Part II. 2019. https://docs.google.com/presentation/d/e/2PACX-1vRU4J_rln00UVD7pNPT0_02NOad0HfSk_UKqRIOv29y8QkMApIEDIyjc0Ot_VE_paV6WBW29Fh_V-iN/pub?start=false&loop=false&delayms=3000#slide=id.g286e6f468f_0_66
- [10] MARC21 Format for Bibliographic data. Library of Congress. Update No. 30 (May 2020) <https://www.loc.gov/marc/bibliographic/>
- [11] Tennant, Roy MARC must die. In: Library Journal, 2010. 05. 21. <https://www.libraryjournal.com/?detailStory=marc-must-die>
- [12] Thomale, Jason. Interpreting MARC: Where’s the Bibliographic Data? Code4Lib Journal, 11. 2010. 09 21. <https://journal.code4lib.org/articles/3832>
- [13] Ungváry Rudolf: A besorolási adatcsere-formátum bővülése. A legutóbbi két évtized fejleményei. In: TMT. 2011. 09. 10., 58. évf. 9. sz., p. 371–386. <https://tmt.omikk.bme.hu/tmt/article/view/868/10649>
- [14] Ungváry Rudolf: Besorolási, szabványosított, normatív vagy „autorizált”. In: TMT, 2019. 06. 24., 66.évf., 6. sz., p. 328–342. <https://tmt.omikk.bme.hu/tmt/article/view/12309/14064>
- [15] Ungváry Rudolf: MARC21 tartalmi adatmezők használata jelentősebb nagykönyvtárakban. Egy elemzés tanulságai. – Networkshop 2020. szeptember 2–4. <https://nws.comp-rend.hu/program/> [online konferencia]