



Még mindig arra a temetésre várunk...

Az új generációs INTERMARC-ban rejlő ígéretek és kihívások

DANCS Szabolcs

Az ismertetés PEYRARD, Sébastien – ROCHE, Mélanie (2018): Still waiting for that funeral: the challenges and promises of a next-gen INTERMARC (elektronikus elérése: <http://library.ifla.org/2204/1/141-peyrard-en.pdf>) írása alapján készült. A cikket a szerzők Creative Commons Attribution 4.0 licenc alatt tették elérhetővé. A licenről bővebben: <http://creativecommons.org/licenses/by/4.0>

A Könyvtári Figyelő jelen számában közölt, Tóvári Judit tollából származó terminológiai vitaindítóhoz, valamint általában a katalogizálás szabályozásának megújításához némiképp kapcsolódik, azt új, gyakorlati szempontokkal gazdagítja a következőkben tárgyalt franciaországi tapasztalat, amely elsősorban a helyi adatsere-formátum korszerűsítésére vonatkozik.

A *Bibliothèque nationale de France* (BnF) által használt MARC-alapú adatsere-formátum, az INTERMARC immár negyven éves múltra tekint vissza. Megújításának közelebbi célja, hogy kompatibilissé tegyék a szemantikus technológiákkal és az új bibliográfiai modellekkel (FRBR, LRM). A francia nemzeti könyvtár metaadat-szakértőinek meggyőződése, hogy a MARC megtartása elősegíti a szakmai folytonosságot és a bibliográfiai szaktudás fejlődését, valamint a nem bibliográfiai felhasználók igényeit is kielégíti. Célkitűzésük a kapcsolt adatok irányába való elmozdulás a MARC rekordok „finomításával”, hordozhatóságuk növelésével. Ennek kontextusa a *Bibliográfiai átállás* nemzeti projektje, amely a katalogizálás LRM-ét (könyvtári referenciamodell alkalmazását) jelenti az RDA implementálásával.

Bevezetés

A kiinduló problémát az jelenti, hogy az információs technológia gyorsabban fejlődik, mint a bibliográfiai formátumok. Az FRBR megjelenése a századfordulón tovább bonyolította a helyzetet. Tíz évvel később az RDA-nak, a szemantikus webes megközelítés az és FRBR közös sarjának megszületése, az IT és a bibliográfiai formátumok nászának végéhez vezetett, miután a szabályzat formátumsemleges katalogizálási szabályokat fektetett le.

Nem kétséges, hogy a MARC egyelőre nem tudott megfelelni az új szabályok és modellek által támasztott kívánalmaknak, de hozzá kell tenni, hogy a „MARC formátumok” alatt a legtöbben magát a MARC21-et értik, miközben a MARC21 hiányosságait nem lehet számon kérni az egyéb MARC-formátumokon. Ezt az álláspontot képviselte legalábbis a francia nemzeti könyvtár (BnF), amikor nekifogott saját formátuma, az INTERMARC „generáljavításához”, amelynek célja, hogy a formátumot a szemantikus webbel, az új bibliográfiai modellekkel és katalogizálási szabályokkal kompatibilissé tegyék. A projekt jelenlegi szakaszában annyit már le lehet szögezni, hogy az új generációs INTERMARC entitásorientált lesz, jelentős mértékben fog támaszkodni a kontrollált szótárakra, és meta-metaadatokat fog tartalmazni.

A MARC21-nek halnia kell?

Immár tizenöt éve annak, hogy a könyvtáros világ megrendülve értesülhetett a MARC halálhíréről. A cikk szerzői ekkoriban még nem voltak könyvtárosok, a MARC születésekor pedig még csak nem is éltek, ezért szakmai elfogultsággal nem vádolhatók. Nézetük szerint a tény, hogy a temetésre még nem került sor, azt sugallja, hogy a haláleset sem történt még meg. A MARC elavultságát emlegető egykori kritikák a MARC21 vonatkozásában helyesnek bizonyulhatnak, de példának okáért az UNIMARC már a kezdetektől épít a lokális azonosítókra és a rekordkapcsolatok lehetőségére. A Kongresszusi Könyvtár 2012-ben döntött az RDA adaptálásáról, azóta a MARC21-et igyekeznek az új szabályokhoz igazítani. Az adatmodellt érintő strukturális revízió azonban egyelőre elmaradt, míg az UNIMARC esetén ez már 2010-ben lezajlott a FRAD-nak való megfelelés érdekében a besorolási adatok kezelése vonatkozásában, valamint az elmúlt pár évben a bibliográfiai formátumot illetően is, amely immár

magában foglalja a *Mű és Kifejezési forma* FRBR-entitásokat is. A BnF az UNIMARC által kijelölt utat kívánja tehát folytatni.

A nemzeti bibliográfiai ügynökség érdeke

A MARC-kal szemben hangoztatott egyik érv, hogy a MARC-rekordok olvasására csak néhány „csúcs katalogizáló” képes. A BnF-ben mintegy háromszázan kezelnek MARC-adatokat napi szinten, nem beszélve a metaadat-szakértőkről és oktatókról. A francia katalogizálók a számítógépes katalogizálás kezdetétől (80-as évek vége) közelről ismerik a MARC-formátumokat, és jóval távolabb állnak tőlük az angol nyelven nyugvó egyéb formátumok, mint például az XML, amitől egyesek a MARC leváltását várták.

A BnF húsz éve egy saját fejlesztésű MARC-alapú katalogizáló rendszert használ, és több mint negyven éve az INTERMARC formátumot, amely a MARC21 és UNIMARC sajátos elegye. Az általános felépítést a MARC21-től, míg egyes alapvető sajátosságokat (mint például a besorolási [*autorizált?*] és bibliográfiai rekordok közötti kapcsolódási lehetőségek) az UNIMARC-tól örökölte, ugyanakkor mindkét formátumnál részletesebb, kidolgozottabb. Húsz évvel ezelőtt komoly változáson, konszolidáción ment át a formátum, hogy megfeleljen a különböző dokumentumtípusoknak és gyűjteményi elvárásoknak. Ma elsődleges szempont az interoperabilitás, ugyanakkor ezúttal is alapkérdés, hogy nulláról kezdjék-e a fejlesztést vagy alapozzák egy olyan megoldásra, amelyet immár jól ismernek és kezelnek. A nemzeti könyvtár úgy döntött, nem kockáztat egy új formátum bevezetésével, a nemzeti bibliográfia előállítójaként egy ilyen mérvű, nehezen menedzselhető változást nem is engedhet meg magának, mivel komoly fennakadáshoz vezethetne a kötelempéldányok feldolgozásában. Ugyanezt diktálta a francia könyvtárügyi többi szereplője iránti felelősségérzet, hiszen a könyvtárak Franciaországban továbbra is MARC-ot használnak, és a belátható jövőig maradnak is ennél a formátumnál.

Az új generációs INTERMARC felé

A BnF szakemberei nincsenek meggyőződve arról, hogy a BIBFRAME fejlesztése a MARC-alapú struktúra leváltásához vezet, mindenképp az IFLA-LRM modell interpretálása okoz problémát. Elismerik, hogy az RDF-ben való publikálás kulcsfontosságú a szemantikus kereshetővé tétel és újrafelhasználás

szempontjából, ugyanakkor úgy gondolják, fontos éles különbséget tenni az adat létrehozásának módja és azon mód között, ahogy az adatot tároljuk és publikáljuk. Az RDF-tripletek tárolása és publikálása nem szükségszerűen jelenti azt, hogy a katalogizálás során hozzuk létre őket.

Az alkalmazott módszer

A francia nemzeti könyvtár munkatársai mindele előtt körüljárták, szükséges-e a MARC-alapú formátum megőrzése, illetve egy új generációs INTERMARC kifejlesztése, és mennyiben áll ez összhangban a bibliográfiai adatok újrafelhasználásának céljával, a kapcsolt adatokra vonatkozó alapelvekkel. Úgy találták, hogy a belső formátum megőrzése egyfelől nem jelent forradalmi változást a katalogizáló munkatársak számára, másfelől teljes mértékben harmonizál azokkal a célokkal, amelyeket a *Bibliográfiai átmenet* (<https://www.transition-bibliographique.fr>) programja képvisel.

A legfontosabb kérdéseket egy kétnapos rendezvény (World Café) keretében, párhuzamosan futó műhelymunkák keretében tisztázták, gondosan ügyelve arra, hogy elvont eredmények helyett kézzelfogható, konkrét megoldási javaslatok szülessenek a formátum átalakítása tárgyában. A megvitatott témák:

- Entitások és kapcsolataik: hogyan tudja a MARC hatékonyan kifejezni az FRBR-t?
- Meta-metaadatok: hogyan tudunk hatékonyan metaadatokat szolgáltatni a metaadatainkról úgy, hogy a felhasználók és metaadat-menedzserek megértsék, milyen minőséggel rendelkeznek az adott metaadatok egy bizonyos előállítási folyamat kontextusában?
- Hivatkozási információ megoldása: olyan környezetben, ahol a géppel olvashatóság és az interoperabilitás tekintetében a szemantikus web szempontjai a mérvadók. Hogyan tehető a MARC teljes mértékben kapcsolattadat-kompatibilissá?
- Szintaxis és formátumkonverzió: a jelenlegi szintaxis (ISO-2709) továbbra is relevánsnak tekinthető? Milyen változás vagy egyszerűsödés várható a formátumkezelésben, illetve miként hat mindez a futó konverziókra, amelyek célja, hogy a BnF adatai nemzetközi szinten támogassák az interoperabilitást?

A workshopok sikeresnek bizonyultak, lefektették azokat az alapelveket, amelyeket követve az INTERMARC továbbfejleszhető. Ezek után mun-

kacsoportokat hoztak létre, amelyek többek között beazonosították a szükséges entitásokat, entitás-kapcsolatokat és ismérveket, valamint elemezték az INERMARC aktuális állapotát: milyen mezőket és almezőket tartsanak meg, definiáljanak újra, hozzanak létre. A 2017 májusa és decembere között tartott megbeszéléseket újabb találkozók követték, ezúttal az IT személyzet és a Metaadat-osztály között, hogy megállapítsák, mennyiben ültethetők át gyakorlatra az elméleti előkészítés konklúziói. Ekkor megérett az idő, hogy a Metaadat Ügyviteli Bizottság véleményezze az újgenerációs INTERMARC-ra vonatkozó alapelveket, amelyre 2018 februárjában került sor. Az ezt követően megfogalmazott Magna Carta a főbb alapelveken kívül útmutatásokat tartalmazott azon szakemberek számára, akiknek a feladata lesz a „jó öreg” INTERMARC megújítása.

Az új formátum funkciói

A Magna Cartában lefektetett három legfontosabb alapelv: az entitás-kapcsolatok koncepciójának implementálása az IFLA LRM-ben lefektetett modell alapján; a kapcsolatok kifejezése kifinomultabb módszer alkalmazásával, mint amilyenre a „hagyományos rekordok” lehetőséget biztosítanak; a kontrollált értékek kifejezése egyedi entitásokon keresztül.

A katalogizálók már jóval a web megszületése előtt kapcsolt adatokban gondolkodtak, hiszen a világ dolgait már régóta dokumentumokkal (rekordokkal) írják le, illetve azonosítókkal kapcsolják össze. Az újgenerációs formátum (INTERMARC-NG) célja, hogy az FRBR-ben meghatározott elvont magentitásokat összhangba hozza a róluk készült leírásokkal. Az ez alapján megfogalmazott alapelvek a következők:

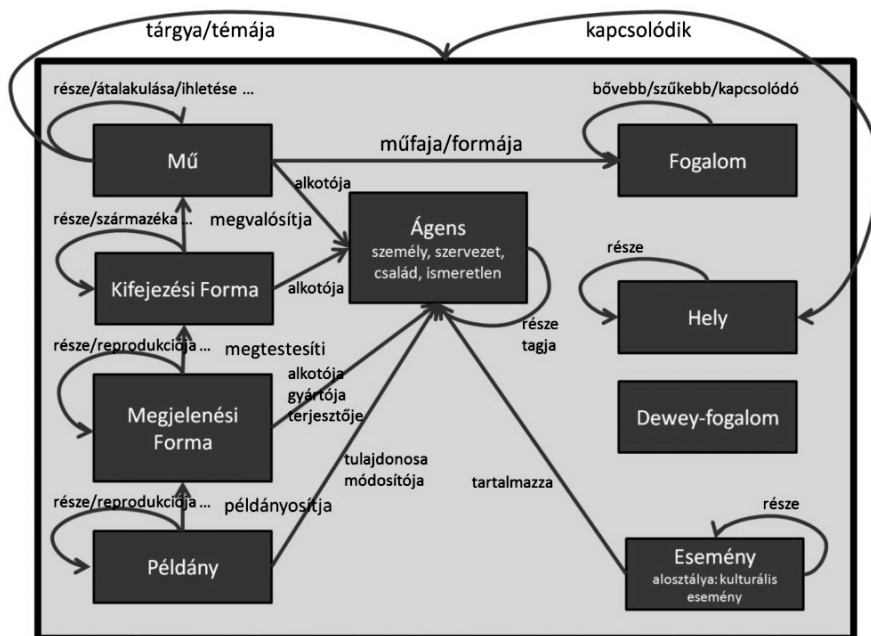
- Minden, az INTERMARC-NG szerint leírt entitásnak külön rekorddal kell rendelkeznie.
- Az INTERMARC-NG szerinti rekordok egyetlen adott entitást írnak le, amiből következik az egyes rekordokban jelenleg elegyedő információk szétválasztásának igénye, hiszen egy MARC-rekord tipikusan tartalmaz információt a Műre, a Kifejezési formára és a Megjelenési formára vonatkozóan. A Mű és a Kifejezési forma szintjeire vonatkozó elemeket eltávolítják a bibliográfiai rekordból, és külön Mű- és Kifejezési forma rekordokat hoznak létre, amelynek az eredménye egy „megsoványodott” Megjelenési forma rekord. A Kifejezési forma rekordok létrehozásakor nulláról indulnak, míg a Mű re-

kordok esetén az egységesített címek besorolási állományára tudnak építeni, új ismérvekkel és kapcsolatokkal bővítve a rekordokat.

- Az entitások közötti kapcsolatokat az adott entitásokat leíró rekordok kapcsolódásaival szükséges kifejezni. A Megjelenési forma és a Kifejezési forma közötti kapcsolatot a Megjelenési forma rekordjából a Kifejezési forma rekordjára irányuló (illetve mutató) kapcsolattal szükséges kifejezni, az utóbbi azonosítójának alkalmazásával. Ugyanaz a mechanizmus alkalmazandó mindenhol, jóllehet jelenleg különböző az eljárás a bibliográfiai és besorolási (autorizált)

formátumok esetén.

A BnF munkatársai mindamellett úgy döntöttek, hogy a Nomen (LRM-E9) és az Időtartam (LRM-E11) entitásokat nem külön rekordokban, hanem ismérvekként kezelik, vagyis a Nomen és Időtartam entitáskapcsolatai közvetlenül azoknak az entitásoknak a rekordjaiban fejeződnek ki, amelyekhez kapcsolódnak, a kapcsolatot pedig az őket tartalmazó mező vagy almező minősíti. Létrehozhatók továbbá a Res entitás alosztályai olyan speciális entitások kifejezésére, mint amilyenek a Fiktív ágens, Fiktív hely, Fogalom, Esemény, Kulturális Esemény, Dewey-fogalom, kontrollált értékek.



1. ábra

Az INERMARC-NG egyszerűsített adatmodellje, entitásonként egy rekorddal és kapcsolatonként egy adatmezővel

Zajlik a tartalomfeltáráshoz használt indexelő nyelv (Rameau) újragondolása. Kihívást jelentenek az adott mű kontextusában jelentőségüket elnyerő entitáskapcsolatok: miközben pl. a főzés önmagában nem kapcsolódik Franciaországhoz vagy a 19. századhoz, a fogalom, hely és időtartam entitások közötti kapcsolat igazzá válhat egy a 19. századi francia konyháról szóló Mű leírása esetén. Ennél fogva úgy definiálják át a téma jelölésére szánt adatmező almezőit, hogy azok a tárgyszavak struktúrájának tükrözése helyett azt fejezzék ki, ahogy a Mű témája az adott, egymáshoz kapcsolódó vagy egymással csoportot alkotó entitások révén kifejeződik.

Meta-metaadatok

A BnF bibliográfiai adatai 2014 óta újrafelhasználhatók, így például a forrás megadásával szabadon integrálhatók külső adatbázisokba. A Wikidata például a BnF vonatkozó rekordjának megfelelő adatmezőjéből veszi a zeneszerző nyelvére vonatkozó információt. Jó volna, ha nemcsak a rekordok, de az egyedi információk tekintetében is jelölhető lenne az adatforrás a rekordokban, ezzel is jelezve az adatok megbízhatóságát. Ennek érdekében négy meta-metaadat mezőt hoztak létre, ezek a következőkre vonatkoznak:

- a Megjelenési forma leírásának forrása;

- ugrópont (link) belső forráshoz (ha a BnF katalógusában megtalálható a Megjelenési formára vonatkozó leírás);
- ugrópont (link) külső webes forráshoz URL-en keresztül;
- külső webes forrás esetén az URL lekérdezésének dátuma.

A Nomen entitás (LRM-E9) ismérveit szintén meta-metaadat elemekként definiálják, külön almezőket rendelve hozzájuk olyan információk rögzítésére, mint a kódolási séma, a célközönség, a felhasználási környezet, a nyelv, írásrendszer. Ezek a meta-metaadatok jóformán minden MARC-mező esetén relevánsak, ezért egységesen ugyanazt a mezőkódot rendelték hozzájuk, bármely mezőn belül használják is őket. A meta-metaadatok almezőkódjai a többitől eltérően nagybetűsök. Ezekon kívül további meta-metaadatokat használnak a rekordtörténet, adott adatmező eredete, a metaadatokra vonatkozó licencinformációk feltüntetésére.

Kontrollált szótárak

A szemantikus webnek való megfelelés azt jelenti, hogy adatainkat a gépek képesek elemezni, vagyis minden adatunkat egy egyedi értékekből felépülő kontrollált szótár részévé kell tenni, hogy az alábbi ismérveket rendelkezésükre hozzák:

- preferált címke és címkeváltozatok;
- kapcsolat (link) a szűkebb vagy tágabb fogalmakat lefedő, illetve egyéb módon kapcsolódó értékekhez;
- a felhasználásra és történetre vonatkozó megjegyzések;

- kód, ha létezik szabványos kód (pl. ISO nyelv-kódok);
- az értékhalmoz (kontrollált szótár), amelyhez az adat tartozik.

Ezek az információk az INTERMARC-ban már most kifejezhetők. A továbbfejlesztés keretében az egyes kontrollált értékek eltérő entitásokként értelmezendők, így külön rekordokat képeznek, illetve külön azonosítóval látják el őket.

Hogyan hatnak mindezek a formátum általános szintaxisára? A meta-metaadatok nem fejezhetők ki fix hosszúságú adatmezők, illetve indikátorok esetén, a kontrollált értékeket pedig ugyanezekben az esetekben csak kóddal lehet rögzíteni, azonosítóval nem. A francia kollégák ezért úgy döntöttek, elhagyják a fix hosszúságú mezőket és indikátorokat, ami nem befolyásolta a formátumszintaxis homogenitását.

Konklúzió

A MARC formátum halálát jósló elemzések nem terjeszthetők ki minden MARC-alapú formátumra. A BnF munkatársai ezt felismerve a formátum temetése helyett a felélesztésének lehetőségeit kezdték el tanulmányozni. Az újgenerációs INTERMARC definiálása a szervezet szempontjából különös kihívást jelentett, ugyanakkor a szakértők egyetértettek a főbb alapelvekben. Az ismert cikk a fejlesztés korai fázisában született, a formátum teljes kidolgozása, a gyakorlatba való átültetése az IT személyzet, majd a katalógizálók által, egy későbbi írás témája lehet.

Beérkezett: 2018. július 27.

Hanák Gábor az idei Kosáry Domokos-díjas

A Kosáry Domokos Könyvtár és Levéltár kezdeményezésére a Szent István Egyetem Szenátusa 2013-ban Kosáry Domokos-díjat alapított a könyvtár névadója tiszteletére születésének 100. évfordulója alkalmából. Az elismerés a könyvtárügy területén, a könyvtári, levéltári munka szervezésében elért kimagasló eredmény, az egyetemtörténeti, agrártörténeti kutatások terén végzett magas szintű tevékenység, illetve a könyvtár, a könyvtárügy támogatásának elismeréseként adományozható.

Ebben az évben Hanák Gábor történész, dokumentumfilm-rendező, riporter, az OSZK Történeti Interjúk Tárának c. igazgatója, jelenleg az általa alapított Cseh Tamás Archivum tanácsadója vette át a Szent István Egyetem Szenátusának tanévzáró ünnepi ülésén 2018. június 15-én az egyetem rektorától és a könyvtár mb. vezetőjétől. A kitüntetést magas szintű, több évtizeden át végzett szakmai és vezetői tevékenységének elismeréseként, a Szent István Egyetem Kosáry Domokos Könyvtár és Levéltár törekvéseinek, történeti kutatásainak példaértékű támogatásáért, történészként és közgyűjteményi szakemberként a felsőoktatási könyvtárak, egyben az egész magyar könyvtárügy támogatásáért kapta a díjazott.

Forrás: Részlet Koósne Török Erzsébet (a SZIE Kosáry Domokos Könyvtár és Levéltár ny. főigazgatója) és Katona Adrienn a Kosáry Domokos-díj Kuratórium titkára tájékoztatójából