


Egy fogalom diadalútja (?)

Entitások a könyvtár-informatikában és a könyvtárosképzésben

A concept's pathway to success (?)

Entities in library informatics and LIS education

Hubay Miklós Péter 
gvm206@gmail.com

humáninformatikai csoportvezető,
MNMKK Petőfi Irodalmi Múzeum

adjunktus, Eszterházy Károly Katolikus
Egyetem, Digitális Technológia
Intézet, Humaninformatika Tanszék

Beérkezett: 2024.07.15.

Elfogadva: 2024.10.16.

Publikálva: 2024.11.06.

Az entitás fogalma és a bibliográfiai univerzum új szempontú felfogása az 1990-es évek második felében jelent meg a könyvtártudományban. Az azóta eltelt csaknem huszonöt évben az új irányzat egyre nagyobb teret követelt magának, eleinte az elméleti publikációk, később már a gyakorlati implementációk területén is. Ma már nem csupán fogalmi modell (LRM), hanem forrásleírási keretrendszer (RDA), illetve az új típusú metaadat-leírásokat hordozni képes informatikai technológia (RDF) is rendelkezésre áll, az esettanulmányok száma is öröndetesen növekszik. Maga az entitásfogalom azonban – a szakirodalom tanúsága alapján – nem rendelkezik egyetlen, széles körben elfogadott definícióval, amely a könyvtári szakemberek közötti párbeszédet, illetve az új kollégák képzését is megnehezíti. Egyes szerzők bibliográfiai egységként, mások egyedi létezőként, megint mások elemeket tömörítő halmazként hivatkoznak rá. Írásomban e nehézkes heterogenitás mérséklése céljából az entitás kategóriaként történő értelmezésének népszerűsítésére teszek kísérletet, s bemutatom, hogy ennek a megközelítésnek milyen előnyei mutatkoznak a közgyűjteményi informatika, valamint a könyvtárosképzés területén.

entitás, könyvtári katalogizálás, könyvtárosképzés, szemantikus elemkészlet, metaadat

In the second half of the 90s, a new trend appeared in the library and information science: it is about thinking about the bibliographic universe as a network of entities and their connections. Today there are many publications on this topic, however, according to the literature, the concept of entity has not only one but at least three definitions, that hardens the discourse between the experts and the library and information science education. There are authors who define entities as if they were equal to bibliographic resources, others think that they are unique and discrete things. And some of the authors think that the entities are sets or categories of unique things. In this study, in order to mitigate this heterogeneity, I try to strengthen the category view, and I examine the problem from the viewpoint of the metadata librarians and also of the LIS educators.

entity, library cataloging, LIS education, semantic element set, metadata

Cite as / Így hivatkozd: Hubay, M. (2024) Egy fogalom diadalútja (?) Entitások a könyvtár-informatikában és a könyvtárosképzésben / A concept's pathway to success (?) Entities in library informatics and LIS education, Central European Library and Information Science Review (CELISR), 1(3), p. 225–234.

<https://doi.org/10.3311/celistr.37885>



1. Bevezetés

Az entitás fogalmának könyvtári területen történő megjelenését a legtöbb szakember az 1990-es évek végéhez, a Bibliográfiai tételek funkcionális követelményei (*Functional Requirements of Bibliographic Records, FRBR*) néven ismert fogalmi modell megjelenéséhez köti. Ez a tanulmány új zemszögből, a felhasználók által megfogalmazott keresőkérdéseket szem előtt tartva vizsgálja, hogy pontosan miről szól, milyen lényeges tematikai egységekre osztható egy bibliográfiai leírás. Az FRBR kidolgozói úgy találták, hogy a metaadat-leírásoknak minden esetben vannak olyan adatelemei, amelyek túlmutatnak a kézbe vett bibliográfiai forráson, és annál magasabb absztrakciós szinteket is megjelenítenek. A modellben definiált három entitáscsoport, különösen az első, amely a bibliográfiai források lehetséges metaadatelemeit a mű, kifejezési forma, megjelenési forma és példány entitások között osztja szét, az évek során önálló katalogizálási irányzat alapjává vált, számottevő mennyiségű elméleti szakirodalommal és egyre növekvő számú gyakorlati vonatkozással.

Az entitás fogalma azonban zavarba ejtően sokszínűen jelenik meg az egyes szerzők munkásságában, sőt az is előfordul, hogy egy életművön vagy egyazon publikáción belül is többféle értelmezéssel találkozunk. A jelen tanulmány előkészítése céljából végzett, a könyvtartudomány területére koncentrált (tehát a történeti, filozófiai stb. munkákat mellőző) irodalomkutatás eredményei azt mutatják, hogy a definíciós kísérletek három nagy csoportba sorolhatók. Az első szerint az entitás a bibliográfiai egységnek, azaz az egyes metaadat-leírások tárgyának felel meg. A második elgondolás képviselői úgy tartják, hogy az entitások olyan diszkrét individuumokat jelölnek, amelyekről egyedi metaadat-leírások készíthetők, míg a harmadik nézőpont hívei az entitásokat nem konkrét létezőkkel, hanem az azokat befoglaló csoportokkal, kategóriákkal azonosítják. Írásomban elsőként ezeket a nézeteket mutatom be részletesebben.

2. Az entitás fogalmának értelmezései

Az *első* definíciós irányzat, amely szerint az entitás a forrásleírás tárgyául szolgáló bibliográfiai egységként értelmezhető, Svenonius (1992) írásából vezethető le. A szerző a bibliográfiai leírás szabatos meghatározását keresve megállapítja, hogy az minden esetben egy bibliográfiai entitás szurrogátuma, így tehát bármely ilyen entitás leírása (*description*) bibliográfiai leírást (*record*) alkot, amely az adott entitás valamilyen célból hasznos ismerveit tartalmazza. A cikk írója ugyanakkor nem határozza meg egyértelműen, hogy mit kell bibliográfiai entitáson érteni: dolgozata az 1990-ben, Stockholmban rendezett Seminar On Bibliographic Records (Bibliográfiai Rekordok Szeminárium) alkalmából készült, amely pont ebben a kérdésben volt hivatott állást foglalni. A szerző által leírt gondolatokból azonban logikusan következik az, hogy amennyiben a leírás tárgya az adott bibliográfiai egység mint egész, akkor ez az egység jelenti a bibliográfiai entitást is. Ugyanakkor a leírás tárgyai olyan entitások is lehetnek, amelyek több fizikai példány közös jellemzőinek megállapításával határozhatók meg: az ilyen entitások közé sorolja Svenonius (1992) például a kiadást, a művet, sőt a szuperművet is. Baker, Coyle and Petiya (2014) úgy véli, hogy a metaadat-leírások általában egyetlen entitást mutatnak be: az adott bibliográfiai forrást (*resource*), amelynek saját, egyedi tulajdonságai vannak. A szerzők kijelentése a weben található, például Dublin Core segítségével megfogalmazott bibliográfiai leírásokra vonatkozik, de általánosítható a hagyományos, analóg katalógusok tételeinek jelentős részére is.

„Az FRBR kidolgozói úgy találták, hogy a metaadat-leírásoknak minden esetben vannak olyan adatelemei, amelyek túlmutatnak a kézbe vett bibliográfiai forráson, és annál magasabb absztrakciós szinteket is megjelenítenek.”

A fogalom *második* értelmezésének hívei szerint entitásnak tekintünk minden egyes, a világon absztrakt vagy kézzelfogható formában létező dolgot. Ezt a definíciót – időrendben az elsőt – Peter Chen alkalmazta munkájában (1976). Az általa kidolgozott módszertannal, amelynek megnevezését az informatikai területen *egyed-kapcsolat* modellezésnek, míg a könyvtár- és információtudományban leginkább *entitás-kapcsolat* modellezésnek fordítják, alapvetően a relációs adatbázisok elméleti tervezését kívánta meg támogatni. Az FRBR megjelenése óta szakmánkban entitásként értelmezhetjük például azokat a személyeket, testületeket, rendezvényeket, földrajzi helyeket stb., amelyek autorizált névalakjai hagyományosan a bibliográfiai leírások hozzáférési pontjait adják; ugyanakkor az FRBR-ben foglaltak alapján entitásként azonosítjuk az egyes műveket, illetve azok kifejezési és megjelenési formáit is. Az entitások jellemzőit a Nemzetközi Katalógizálási Alapelvekben (*International Cataloging Principles, ICP*) megfogalmazottak szerint annak alapján kell megállapítani, ahogyan az entitás saját magát leírja (IFLA, 2016). Minden egyes entitást egyetlen, kitüntetett névalakkal kell reprezentálni (Tillett 2004), az azonos entításra vonatkozó névváltozatokat összegyűjtve, egymással összekapcsolva kell megjeleníteni (Possemato 2018), a közgyűteményi adatbázisok átjárhatósága érdekében pedig az azonos entitást reprezentáló metaadat-leírásokat nemzeti, illetve nemzetközi szinten is össze kell kapcsolni (Library of Congress, 2008). A definíció értelmében az egyes intézmények által a gyűjteménykezelő rendszerekben épített authority-állományokat vagy ezek világhálón közzétett, korszerűbb megfelelőit, a szótárkódolási sémákat (mint a VIAF, a Geonames vagy a Magyar Nemzeti Névtér) az *entitások* adatgazdag leírásainak halmazaként is értelmezhetjük. Az ezekben a leírásokban közölt információk segítségünkre lehetnek az FRBR-ben leírt egyik felhasználói feladat, avagy használói igény kielégítése során. Ez az *azonosítás* igénye, melynek során az információt kereső használónak el kell tudnia dönteni, azon entitás leírását találta-e meg, amelyet keresett, illetve meg kell tudnia különböztetni egymástól a hasonló jellemzőkkel rendelkező entitásokat (IFLA 1998).

A *harmadik* meghatározás szerint az entitásokat legjobban fogalmi kategóriaként ragadhatjuk meg. A Könyvtári Referenciamodell (*Library Reference Model, LRM* – Riva, Le Boeuf, Žumer 2017) kidolgozói szerint az entitások a fogalmi objektumok elvont osztályait jelentik, az individuumszemlélettel ellentétben tehát az entitás ebben a meghatározásban olyan csoportot vagy halmazt jelöl, amelynek sok különböző eleme, előfordulása (*instance*) – az FRBR magyar fordításában *esete* – lehet. Ezeket az előfordulásokat írják le a bibliográfiai, állomány- és authority adatok. Svenonius (1992) és Taniguchi (2003) is halmazként definiálta az entitásokat, és úgy vélték, hogy a metaadat-leírások minden esetben az adott entitás (vagy a vegyes tartalmú MARC-rekordok esetében több entitás) egy-egy konkrét előfordulásának adatait tartalmazzák. Zhu et al. (2023) a Wikidatában szereplő elemekről szólva megállapították, hogy azoknak egy egyértelműen azonosítható fogalmi vagy materiális entitás egy előfordulását kell reprezentálniuk. Az ICP végén olvasható szöveget a kapcsolat fogalmát az entitások *vagy azok előfordulásai* között meghatározott összeköttetésként definiálja (IFLA 2016, p. 15).

Külön csoportot képviselnek azon szerzők – mások mellett Maxwell (2008) és Jin, Hahn and Croll (2016) –, akik az individuumszemlélet mellett teszik le a voksot, az egyedi entitásokat mégis valamilyen, alkalmasint számos elemet tartalmazó nagyobb egységbe foglalják. Jin és szerzőtársai például a művet, kifejezési formát, megjelenési formát és példányt entitás *halmaznak* nevezik, jóllehet ezt a fogalmat maga az FRBR egyáltalán nem használja, hanem minden esetben

„...az FRBR-ben foglaltak alapján entitásként azonosítjuk az egyes műveket, illetve azok kifejezési és megjelenési formáit is.”



egyszerűen entitásokról, illetve az azokat befoglaló entitáscsoportokról beszél. Ungváry (2010) a diszkrét individuumokat – a szerző szóhasználatában egyszerűen: értékeket – magába foglaló entitástípusokat említi, ezeket a típusokat azonban ő is az FRBR-ben és a FRANAR-ban leírt kategóriákkal azonosítja. Az értelmezést nehezíti ugyanakkor, hogy az entitástípus – sőt két helyen entitásfajta – fogalma az FRBR-ben a fentitől eltérő meghatározással szerepel, s csupán az entitás *előfordulásainak* kategóriáit, nem magukat az entitásokat jelenti: a megjelenési forma entitástípusai ennek megfelelően például a mikromásolat, a reprint vagy a faksimile (IFLA 1998). Az entitástípus fogalmát az ICP is használja egy alkalommal, amikor az integráció alapelveinek magyarázatát adja: a kontrollált névformáknak bármely *típusú* entitás esetében egységes szabályrendszer szerint kell elkészülniük (IFLA 2016, p. 5).

Úgy gondolom, a rendelkezésre álló entitásdefiníciók közül több szempontból is az imént tárgyalt, harmadik meghatározás a legcélravezetőbb, jóllehet, ez a megközelítés, mint látható, jelentősen eltér az entitás fogalmának Peter Chen által adott, eredeti értelmezésétől (1976). Írásom további részében ezért arra teszek kísérletet, hogy a kategóriaszemlélet hasznosságát két példa felhasználásával is igazoljam. A következőkben a közgyűjteményi informatika talán egyik legizgalmasabb irányzatával, a szemantikusweb-technológia alkalmazásával foglalkozom, és megvizsgálom az entitások reprezentációját az ún. szemantikus elemkészletben, abban az informatikai konstrukcióban, amely a jövő közgyűjteményi metaadat-leírásainak alapkövét jelentheti. A tanulmány harmadik fejezete pedig azt mutatja be, hogyan segítheti a könyvtároshallgatók munkáját az, ha az entitások kategóriaszemléletét már a képzés során is tudatosan és következetesen alkalmazzuk.

3. Entitás a szemantikusweb-technológiában

A bibliográfiai leírások és authority adatok számítógépes környezetben való tárolásához elengedhetetlenek a strukturáltságot, illetve a visszakereshetőséget biztosító metaadatkészletek. Ezek egyik csoportját az ISO 2709-es szabvány (2008) szerkezeti követelményeire épülő rendszerek alkotják. Ez a formátum a tárolni kívánt információt *rekordba* szervezi, amelyen belül az egyes adatelemeket hívójelek, illetve almezők használatával lehet kifejezni, amit indikátorértékek pontosíthatnak. Ebbe a kategóriába tartozik a világon használt valamennyi MARC-hívójelkészlet, de ide sorolható a főként német nyelvterületen elterjedt PICA3 és PICA+, illetve a ma már aktív használatban nem lévő MAB is.

Az 1990-es évek második felétől egyre nagyobb teret kaptak azok a metaadat-rendszerek, amelyek a kiterjeszhető jelölőnyelv (Extensible Markup Language, XML) sémáinak segítségével definiálható címkékre épülnek. Az XML-sémák segítségével előállított metaadatleírás-halmazok általánosan alkalmazott informatikai technológiájuknak köszönhetően könnyebben feldolgozhatók és megjeleníthetők, egyszerűbben egymásba alakíthatók, valamint általános hálózati adatátviteli protokollok segítségével könnyebben továbbíthatók egyik helyről a másikra.

A metaadatelem-készletek harmadik, jelenleg ismert legkorszerűbb formája *ontológia*, *szótár* avagy *szemantikus elemkészlet* néven ismeretes, amely valamilyen, a forrásleíró keretrendszer (Resource Description Framework, RDF) három elemre épülő logikáját támogató szintaxisban kifejezett szöveges állományként testesül meg. A szótárakban a metaadatelemeket avagy *tulajdonságokat* nem hívójelek, nem is sémában definiált címkék, hanem egyedi, ám egységesen http-szerkezetű azonosítók, ún. forrásazonosítók (Uniform Resource Identifier,

„A metaadatelem-készletek harmadik, jelenleg ismert legkorszerűbb formája ontológia, szótár avagy szemantikus elemkészlet néven ismeretes...”

URI) reprezentálják. A metaadatelem definíciója az XML-hez képest is jóval specifikusabban rögzíthető: tartalmazza az elem más elemekhez való logikai viszonyait és hierarchikus pozícióját (például: a *barátja* vagy a *munkatársa* reláció az *ismeri* reláció alárendeltje, a *menyasszonya* és a *vőlegénye* relációk egymás inverzei stb.), illetve egyéb tulajdonságait, például értelmezési tartományát és értékkészletét is. A szótárak elemeinek használatával felépülő metaadatleírás-halmazokat már nem rekordnak, hanem gráfnak vagy tudásgráfnak nevezzük akkor is, ha az egyetlen bibliográfiai forrás adatait tárolja, és akkor is, ha egy egész könyvtárnyi forrás leírását foglalja magába.

A szemantikus technológia segítségével definiált metaadatelem-készletek nem csupán adatgazdagságukban nyújtanak többet az ISO 2709-, illetve az XML-alapú készletekhez képest (2008). Ugyanis, amíg előbbieknél jól meghatározott, túl nem léphető keretei vannak, addig a szótárak igény szerint kombinálhatók egymással, így összetett alkalmazások a legegyszerűbb igényeket is kielégítheti. Nem kell aggódni tehát, ha egyetlen szótár relációi nem elegendőek bizonyos forrástípusok leírásához, a probléma könnyen orvosolható, ha az alapvetően egy szótár felhasználásával készült gráfokat más elemkészletek relációival is kiegészítjük.

A szótárak fontos feladata, hogy olyan osztályokat határozzanak meg, amelyekbe a gráfban leírni kívánt individuumok besorolhatók. Ha egy ilyen, egyedi URI-val jelölt individuumról, például *Bodriról* vagy *Lassie-ről* kijelentjük, hogy a *kutyák* osztályába tartoznak, ezt követően már alkalmazhatjuk rájuk a *kutyák* osztályának általános tulajdonságait: mi a fajtája, mekkora a marmagassága, milyen hosszú a szőre, milyen elsődleges feladata van stb. Ezen tulajdonságok értelmezési tartománya tehát a *kutyák* osztálya. Amennyiben elfogadjuk az *entitás* fentebb javasolt, kategóriaként történő meghatározását, akkor azt mondhatjuk, hogy a szótárban definiált minden osztály egy-egy entitásnak felel meg, az osztály elemeit pedig a konkrét entitás-előfordulások adják. Ezen osztályok segítségével a tulajdonságok értékkészletét is kifejezhetjük, amely azt mutatja meg, hogy a tulajdonsághoz rendelt konkrét értékek mely osztályokból származhatnak. Így például a művek alkotói és közreműködői kizárólag az *Ágensék* osztályának tagjai közül kerülhetnek ki.

Az osztályt nagyon egyszerűen, a webontológia-nyelv (Web Ontology Language, OWL) egyik relációja segítségével lehet definiálni, amelynek használata során rögzíteni kell az osztály saját URI-ját. Ezt követi az osztály főbb paramétereinek felsorolása, amely a szerző több, közgyűjteményi célú szótár megvizsgálása során szerzett tapasztalatainak alapján döntően az alábbiak rögzítését jelenti:

- az osztály megnevezése, címkéje (*rdfs:label*), akár több nyelven is, amely a képernyőn történő megjelenítést szolgálja;
- az osztály rövid, maximum egy-két mondatos leírása, amelynek kifejezéséhez több szótárból is választható reláció, a leggyakoribbak a *skos:definition*, a *dc:description*, illetve az *rdfs:comment*;
- az osztály hierarchikus elhelyezkedésére, logikai viszonyaira vonatkozó közlések, így például az osztály fölérendeltje (*rdfs:subClassOf*), az osztállyal egyenértékű, más szótárban definiált osztály (*owl:equivalentClass*), vagy az a diszjunkt tartalmú másik osztály, amellyel az adott osztálynak semmiképp sem lehet közös eleme (*owl:disjointWith*);
- adminisztrációs metaadatok, például az osztály létrehozásának, illetve tulajdonságai módosításának időpontja (*dcterms:issued*, *dcterms:modified*);

„A szótárak fontos feladata, hogy olyan osztályokat határozzanak meg, amelyekbe a gráfban leírni kívánt individuumok besorolhatók.”



- különböző megjegyzések: használati megjegyzés (*skos:scopeNote*), az osztály szerkesztőinek megjegyzései (*skos:editorialNote*), az elvégzett módosításokra vonatkozó megjegyzések (*skos:changeNote*, *skos:historyNote*), vagy éppen, indokolt esetben, ha az értelmezést segíti, példa az osztály előfordulásai közül (*skos:example*).

A kialakított és a fent bemutatott paraméterek megadásával jellemzett osztály használatba vétele, azaz a konkrét entitás-előfordulások besorolása ezt követően úgy történik, hogy a tudásgráfban az előfordulást reprezentáló azonosítót összekötjük az osztály (azaz entitás) azonosítójával:

```
<bf:Work rdf:about="http://a besorolandó entitás-előfordulás URI-ja">
```

Tehát a példában említett előfordulás a művek osztályába tartozik. A közgyűjteményi alkalmazású szótárak nagy számban hozzáférhetők a világhálón, elemeik szabadon felhasználhatók az intézmények adatbázisainak konverziója és közzététele során. Közülük kiemelkedik a Kongresszusi Könyvtár és informatikai partnerei által kidolgozott és a MARC adatcsere-formátum utódjának szánt BIBFRAME (*bibliographic framework*, bibliográfiai keretrendszer) elemkészlet, amelynek első, normatív dokumentuma 2012-ben jelent meg (Library of Congress, 2012), a szótár jelen sorok írásakor pedig már a 2.3.0 verzióban hozzáférhető. Sajátos entitásszerkezetének részleteiről a cikk harmadik részében esik bővebben szó, így ide csupán egy nevezéktani érdekességről szóló megjegyzés kívánkozik. A szótárakban általánosan alkalmazott gyakorlat – és a segítségükkel kialakítandó tudásgráfok értelmezését is segíti –, hogy az osztályok/entitások megnevezése nagybetűvel, míg a metaadatelemeké kisbetűvel kezdődik. A BIBFRAME azonban egyedi abban, hogy bizonyos tulajdonságok, illetve az azokkal használható osztályok a kezdőbetűtől eltekintve azonos megnevezést kaptak; megítélés kérdése, hogy ez zavart okoz vagy belső logikát teremt.

```
<bf:media>  
  <bf:Media rdf:about="http://id.loc.gov/vocabulary/mediaTypes/n">  
    <rdfs:label>unmediated</rdfs:label>  
  </bf:Media>  
</bf:media>
```

A rövid gráfrészlet magyarázata: a *media* metaadatelem (kisbetűvel) értékei a BIBFRAME szótárban meghatározott módon, kizárólag a *Media* osztály (nagybetűvel) tagjai közül származhatnak. A fenti esetben *unmediated* (*hordozó nélküli*) a tulajdonság értéke, amelyhez a <http://id.loc.gov/vocabulary/mediaTypes/n> egyedi azonosító tartozik.

4. Oktatási vonatkozások

Az Eszterházy Károly Katolikus Egyetem Digitális Technológia Intézete által koordinált könyvtártudományi alap-, illetve mesterképzésben három olyan tanegység is megtalálható, amelyek tematikájában kisebb-nagyobb helyet kap az entitásokra épülő forrásleírás. Ezek közül hallgatóink legelőször a Metaadat-tárolási technológiák elnevezésű tantárggyal találkoznak, amely az alap- és mesterképzés kínálatában is szerepel (EKKE 2023a, 2023b). Az alapképzés tanterve szerint a tanegység teljesítése az utolsó előtti félévben ajánlott, ekkorra a hallgatók már

„A szótárakban általánosan alkalmazott gyakorlat – és a segítségükkel kialakítandó tudásgráfok értelmezését is segíti –, hogy az osztályok/entitások megnevezése nagybetűvel, míg a metaadatelemeké kisbetűvel kezdődik.”

a könyvtártudomány legtöbb területét, így például az előfeltételül szabott, két féléves, az ISBD rendelkezéseire épülő Bibliográfiai adatfeldolgozás tanegység anyagát is jól ismerik. A Metaadat-tárolási technológiák kurzus több ismeretkört – így például a szemantikus technológia könyvtári alkalmazását – alapozó jelleggel tárgyal, az ismeretek elmélyítése a mesterképzés tantervében megjelenő tárgyak keretein belül történik. A félév első heteiben a metaadat fogalmának körülhatárolásával, a metaadatrendszerek csoportosításával foglalkozunk, majd részleteiben tárgyaljuk a könyvtár-informatikai munka jelenlegi alapját adó MARC- és MARC 21-formátumot és más, ISO 2709-es szerkezetre épülő adatcserezabványokat (2008). A MARC/XML bemutatásán keresztül a kiterjeszhető jelölőnyelvre alapuló adattárolás alapjaival is megismertetjük a hallgatókat.

Ezt követően az entitásalapú forrásleírási irányzat elméletét adó szövegkorpusz, az FRBR és az LRM megismerésével foglalkozunk. Meglátásom szerint itt is kimondottan hasznos, és a későbbi tananyagrészek megértését szolgálja, ha a jelen írás második fejezetében javasoltak szerint az entitásokat fogalmi kategóriának, azaz a konkrét előfordulásokat befoglaló halmaznak tekintjük. E megközelítést alkalmazva ugyanis az *entitás* párhuzamba állítható a hallgatók által ekkor már jól ismert *ISBD-adatcsoport*tal. Mindkettőre igaz, hogy a bibliográfiai forrásokról készített leírások szerkezeti alapegységét adja, illetve hogy saját metaadatelemekkel rendelkezik. Egy ISBD-re, valamint a jelenleg érvényben lévő leírási szabványokra és könyvtári szabályzatokra épülő forrásleírás készítésekor a katalogizáló szakember pontosan tudja, hogy a kezében tartott forrás esetében az adatcsoportok adatelemeit milyen sorrendben, milyen konkrét értékkel kell rögzítenie. Az entitásalapú bibliográfiai leírások esetében a különbség mindössze annyi, hogy a forrásleírás sorvezetőjét már nem az adatcsoportok, hanem az entitások adják. A kérdés tehát az, hogy mit közöl a forrás, és milyen metaadatelemek segítségével írható le, ami a benne rejlő szellemi tartalomra (azaz a műre) vagy a tartalom megtestesítésére használt jelölési rendszer sajátosságaira (azaz a kifejezési formára) vonatkozik, milyen ágensek vettek részt a megjelenési forma elkészítésében és így tovább.

Az entitásokat az egymást követő fogalmi modellek egymástól eltérően szervezik. Az FRBR entitás csoportokat és ún. magentitásokat mutat be, míg a jelenleg „érvényben lévő” Könyvtári Referenciamodell hierarchikus szerkezetet vázol. Úgy vélem, ezen a ponton kifejezetten fontos felhívni a hallgatók figyelmét arra, hogy bizonyos esetekben, jóllehet az entitás megnevezése nem változott, az egyes modellek eltérő meghatározást közölnek például a művek esetében. Ezeket a definíciókat helyezte egymás mellé Seikel és Steele írása (2020), amely a tanegység kötelező olvasmányai között is szerepel. A cikk a két fogalmi, elméleti forrás mellett két gyakorlati implementáció, az RDA bibliográfiai keretrendszer (*Resource Description and Access, Forrásleírás és hozzáférés*), illetve a BIBFRAME szemantikus elemkészlet entitásainak definícióit olvassa össze, amelynek áttekintésével a hallgatók szélesebb körű ismereteket szerezhetnek a bibliográfiai univerzum lehetséges modellezéseivel kapcsolatban, továbbá az egyes definíciók vizsgálata során kialakíthatják saját álláspontjukat, s megfogalmazhatják kritikai észrevételeiket.

A tanegység konzultációin a hallgatók az RDA alapjaival is megismerkednek. A 2000-es évek első felétől lényegében mind a mai napig folyamatosan fejlesztett keretrendszer az entitások mentén szervezett bibliográfiai leírások elkészítéséhez ad előírásokat és javaslatokat (sok esetben opcionális jelleggel), illetve definiálja az entitásokat jellemző, a leírásokban rögzíthető metaadatelemeket, entitásonként akár több százat is. Ezenkívül minden metaadatelemhez, illetve a kötött értékészletű elemek használatkor alkalmazandó, ún. szótárkódolási sémák értékeihez

„Az entitásokat az egymást követő fogalmi modellek egymástól eltérően szervezik.”



egyedi azonosítókat (URI-kat) is rendel, amelyek felhasználhatók a bibliográfiai leírás RDF-logikára épülő tárolása során, de beleírhatók akár a MARC-rekordokba is.

Az RDA-val kapcsolatban a szakma gyakorlatában s ezért a tanórákon is igen gyakran használt fogalom az *alkalmazási profil*. Maga a kifejezés ugyanakkor nem csupán az RDA kontextusában értelmezhető: az alkalmazási profilok dolga kijelölni egy metaadatséma valamilyen konkrét (intézményi, nemzeti, regionális stb.) használatának kereteit: rendelkezhetnek a használandó metaadatelemekről, illetve megkötések tartalmazhatnak azok ismételtetésére, kötelezőségére, értékeinek alakjára, szerkezetére vonatkozóan stb. Az RDA-alkalmazási profil megalkotását az a tény teszi szükségessé, hogy az RDA zavarba ejtően sok metaadatelemet, illetve leírási opciót kínál a forrásleírásokhoz. Ezek közül célszerűen az adott ország bibliográfiai számbavételét koordináló, nemzeti bibliográfiai ügynökségnek – így Magyarországon az OSZK-nak – kell megjelölnie azokat az alkalmazandó szabályokat, továbbá a leírásokban feltüntetendő entitásokat és adatelemeket, amelyeket az egyes forrástípusok RDA szerinti leírásához használhatnak az egyes intézmények, így biztosítva az átjárhatóságot és a leírások egységes szerkezetből következő újrafelhasználhatóságát az aggregációs szolgáltatásokban (Ilácsa 2020).

A másik igen jelentős gyakorlati implementáció, amelyet a tananyag tartalmaz, a BIBFRAME bibliográfiai keretrendszer – amelyet, mint korábban említettem, a Kongresszusi Könyvtár azzal a céllal dolgozott ki, hogy az majd egyszer a MARC helyébe lép, tehát a jövőben az intézmények közötti adatcsere nem rekordok, hanem BIBFRAME-alapú RDF-gráfok segítségével fog megvalósulni. A keretrendszer tárgyalását sajátos adatmodellje bemutatásával kezdjük. A Library of Congress munkatársai ugyanis nem követik az FRBR-ben, illetve az LRM-ben vagy akár az RDA-ban leírt négy entitásra épülő megközelítést, helyette három alapvető entitást határoztak meg, amelyek a mű (Work), előfordulás (Instance) és példány (Item) neveket kapták. A gyakorlati tapasztalatok azonban igazolták, hogy szükséges az adatmodell kiegészítése egy speciális, negyedik entitással, mely a mű feletti absztrakciós szinten helyezkedik el, és a csomópont (Hub) címet viseli, meghatározását tekintve pedig leginkább a Svenonius (1992), Yee (2000) és Carlyle (1996), továbbá Smiraglia (2003) által említett *supermű*höz áll legközelebb. Ilyen csomópontok keletkeznek a MARC-ról BIBFRAME-re történő konverziókor például az egységesített címetek, illetve a sorozatcímetek tartalmazó mezők, továbbá a tárgyi hozzáférési pontok adatmezői esetében. Sorozatoknál tehát a csomópont entitás konkrét előfordulásainak adatelemei a 440-es hívójelű MARC-mezőben rögzített megfelelően alakulnak, és ehhez a csomópont entitáshoz kapcsolódik a sorozat tagjaiban reprezentált művek mindegyike.

A BIBFRAME-mel és entitásaival kapcsolatos mélyebb ismereteket a korábban leírtak szerint mesterképzésen adjuk át a hallgatóknak a Metaadatsémák, ontológiák, felhőtechnológiák tárgy anyagába építve. Görög kutatók munkásságát – Zapounidou (2014, 2019) – felhasználva bemutatjuk a BIBFRAME, illetve az RDA szemantikus elemkészlete közötti megfeleltetés nehézségeit, amelyek az eltérő entitásszerkezetből következnek, továbbá szót ejtünk a BIBFRAME-adatok EDM-re (Europeana Data Model) való átültetésének lehetőségeiről és sajátosságairól is. Ebben a tanegységben már van mód néhány BIBFRAME-re épülő forrásleíró felület, például a MARVA vagy a Sinopia bemutatására, amelyek segítségével a hallgatók gyakorlati tapasztalatokat szerezhetnek az entitásalapú forrásleírással kapcsolatban, s ezzel párhuzamosan a BIBFRAME sajátos adatmodellezését is jobban megérthetik, s összevethetik a fogalmi modellek által javasolt négyelemű alternatívával.

„az alkalmazási profilok dolga kijelölni egy metaadatséma valamilyen konkrét (intézményi, nemzeti, regionális stb.) használatának kereteit...”

5. Összegzés

A könyvtártudományi-informatikai szakirodalom az elmúlt években, évtizedekben, mint jelen írás első fejezetében láttuk, megdöbbentő heterogenitással igyekezett definiálni az entitás fogalmát. A hazai, illetve nemzetközi publikációkban tapasztalható sokszínűség, bár védhető, mégis jelentősen megnehezíti a szakemberek, illetve a tudományterülettel épp csak ismerkedők dolgát. Írásomban az entitások kategóriaként való meghatározása mellett érveltem, és úgy vélem, ez az álláspont oktatási, illetve könyvtár-informatikai szempontból egyaránt szerencsés választás. A kategóriaszemlélet alkalmazása a cikkben említettek szerint lehetővé teszi, hogy az entitást egyfelől a szemantikus elemkészletben definiált osztállyal, másfelől az ISBD adatcsoportjaival rokonítsuk, amely a felvázolt párhuzamosságokon keresztül minden bizonnyal elősegíti az egyébként valóban igen absztrakt fogalom megértését.

Ennek ellenére a fogalomhasználat egységesedéséig valószínűleg még hosszú idő fog eltelni, az eredményesség jelen pillanatban még legalábbis kétséges. Az előttünk álló átmeneti kor lerövidítéséhez azonban feltétlenül szükségesnek tartom, hogy az új terminológia alkalmazására vonatkozó ismereteket és érveket egyaránt meghallgassák és elsajátítsák a jelen és a jövő közgyűjteményi munkatársai.

Felhasznált és ajánlott forrás

Baker, T., Coyle, K., Petiya, S. (2014) *Multi-Entity Models of Resource Description in the Semantic Web: A comparison of FRBR, RDA, and BIBFRAME*, Library Hi Tech, 32(4), p. 565–566. <https://doi.org/10.1108/LHT-08-2014-0081>

Carlyle, A. (1996) *Ordering author and work records: an evaluation of collocation in online catalog displays*, Journal of the Association for Information Science and Technology, 47(7), p. 538–554. [https://doi.org/10.1002/\(SICI\)1097-4571\(199607\)47:7%3C538::AID-ASI6%3E3.0.CO;2-V](https://doi.org/10.1002/(SICI)1097-4571(199607)47:7%3C538::AID-ASI6%3E3.0.CO;2-V)

Chen, P. (1976) *The entity-relationship model – toward a unified view of data*, ACM Transactions on Database Systems, 1(1), p. 9–36. <https://doi.org/10.1145/320434.320440>

EKKE (2023a) *A I-KVT-BA-N-E kódú, Informatikus könyvtáros BA szak mintatanterve*. Elérhető: https://oktatas.uni-eszterhazy.hu/tik_kalauz/tikkalauz2023/units/I-KVT-BA-N-E.PDF (utolsó elérés: 2024.10.13.)

EKKE (2023b) *A I-KIT-M-L-E kódú, Könyvtár- és információtudomány szak mintatanterve*. Elérhető: https://oktatas.uni-eszterhazy.hu/tik_kalauz/tikkalauz2023/units/I-KIT-M-L-E.PDF (utolsó elérés: 2024.10.13.)

IFLA (1998) *A bibliográfiai tételek funkcionális követelményei*. Elérhető: <http://www.ifla.org/files/assets/cataloguing/frbr/frbr-hu.pdf> (utolsó elérés: 2024.10.13.)

IFLA (2016) *Nyilatkozat a nemzetközi katalógizálási alapelvekről (ICP)*. Elérhető: https://www.oszk.hu/sites/default/files/ICP_2016_magyarul_2019_09_27.pdf (utolsó elérés: 2024.07.09.)

Ilácsa Sz. (2020) *Alkalmazásprofilok készítése RDA-hoz: kísérleti projekt alkalmazásprofilok módszertanának kidolgozásához*, Tudományos és Műszaki Tájékoztatás, 67(9), p. 538–555. Elérhető: <https://journals.bme.hu/tmt/article/view/35331/21044> (utolsó elérés: 2024.07.09.)

International Organization for Standardization (2008) *ISO 2709:2008 Information and documentation – Format for information exchange*. ISO [Online]. Elérhető: <https://www.iso.org/standard/74046.html> (Utolsó elérés: 2024.10.14.)



Jin, Q., Hahn, J., Croll, G. (2016) *BIBFRAME Transformation for Enhanced Discovery*, Library Resources & Technical Services, 60(4), p. 223–235. <https://doi.org/10.5860/lrts.60n4.223>

Library of Congress (2008) *On the Record: Report of The Library of Congress Working Group on the Future of Bibliographic Control*. Elérhető: <http://www.loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf> (utolsó elérés: 2024.07.09.)

Library of Congress (2012) *Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services*. Elérhető: <https://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf> (utolsó elérés: 2024.07.06.)

Maxwell, Robert L. (2008) *FRBR – A Guide for the Perplexed*. ALA, Chicago, 2008. ISBN: 978-0-8389-0950-8

Possemato, T. (2018) *How RDA is essential in the reconciliation and conversion processes for quality Linked Data*, J LIS.it, 9(1), p. 48–60. <https://doi.org/10.4403/jlis.it-12447>

Riva, P., Le Boeuf, P., Žumer, M. (2017) *IFLA könyvtári referenciamodell: a bibliográfiai információk elméleti modellje*. Elérhető: https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712-hu.pdf (utolsó elérés: 2024.07.09.)

Seikel, M., Steele, T. (2020) *Comparison of Key Entities Within Bibliographic Models and Implementations: Definitions, Evolution, and Relationships*, Library Resources & Technical Services, 64(2), p. 62–71. <https://doi.org/10.5860/lrts.64n2.62>

Smiraglia, R. P. (2003) *The History of "The Work" in the Modern Catalog*, Cataloging & Classification Quarterly, 35(3-4), p. 553–567. https://doi.org/10.1300/J104v35n03_13

Svenonius, E. (1992) *Bibliographic Entities and Their Uses*, In: Bourne, R. (szerk.): Seminar on Bibliographic Records: proceedings of the seminar held in Stockholm, 15-16 August 1990: and sponsored by the IFLA UBCIM Programme and the IFLA Division of Bibliographic Control, p. 3–18, Saur, München, ISBN: 3598110855

Taniguchi, S. (2003) *A Conceptual Model Giving Primacy to Text-level Bibliographic Entity in Cataloging: A Detailed Discussion*. Elérhető: <http://user.keio.ac.jp/~taniguchi/report200301.pdf> (utolsó elérés: 2024.07.06.)

Tillett, B. (2004) *Authority Control: State of the Art and New Perspectives*, Cataloging & Classification Quarterly, 38 (3-4), p. 23–41. https://doi.org/10.1300/J104v38n03_04

Ungváry, R. (2010) *MARC21/HUNMARC - a besorolási adatok metaadat-formátuma: főbb jellemzők, fejlődés és problémák*, Könyvtári Figyelő, 56(1), p. 9–70. Elérhető: http://epa.niif.hu/00100/00143/00074/pdf/EPA00143_Konyvtari_Figyelo_2010_1_009-070.pdf (utolsó elérés: 2024.07.06.)

Yee, M. M. (2000) *Lubetzky's Work Principle. The Future of Cataloging: Insights from the Lubetzky Symposium*, April 18, 1998, University of California, Los Angeles, 1, p. 72–104. Elérhető: <https://escholarship.org/uc/item/60d2c41c> (utolsó elérés: 2024.07.06.)

Zapounidou, S., Sfikakis, M., Papatheodorou, C. (2014) *Library Data Integration: Towards BIBFRAME Mapping to EDM*. In: Metadata and Semantics Research: 8th Research Conference, MTSR 2014 Karlsruhe, Germany, November 27–29, 2014, Proceedings, p. 262–273.) Springer. https://doi.org/10.1007/978-3-319-13674-5_25

Zapounidou, S., Sfikakis, M., Papatheodorou, C. (2019) *Mapping Derivative Relationships from RDA to BIBFRAME 2*, Cataloging & Classification Quarterly, 57(5), p. 278–308. <https://doi.org/10.1080/01639374.2019.1650152>

Zhu, L., Xu, A., Deng, S., Hang, G., Li, X. (2023) *Entity Management Using Wikidata for Cultural Heritage Information*, Cataloging & Classification Quarterly, 61(1), p. 20–46. <https://doi.org/10.1080/01639374.2023.2188338>