

Új könyvtári fogalmak gyakorló könyvtárosok szemszögéből

Hubay Miklós
Petőfi Irodalmi Múzeum
hubaym@pim.hu
ORCID: 0000-0002-3241-2988

Gyuricza Andrea
Magyar Tudományos Akadémia Könyvtár és Információs Központ
gyuricza.andrea@konyvtar.mta.hu
ORCID: 0000-0003-3960-5499

As the landscape of libraries changes nowadays, there are more and more new concepts, which are often defined with other, incomprehensible terms in scholarly context. A new trend of library cataloging has appeared, that broke up with the 50-years-old traditions, and besides the simple registering of bibliographic data, it is necessary to put these data in a broader context, in a network. In order to accomplish this task, some of the experimenters use the good old solutions, while others define new theoretical borders. The definitions in the library journals sound very well, while the practice-centered approach is missing.

Now we clear the meaning of some of these library and IT science terms (LRM, discovery, semantic web, MARC21, FRBR, RDA, entity, BIBFRAME, URI, schema.org etc.), so the patron service librarians and also the catalogers can understand the transformation of LIS, can interpret the connections, correlations, and the main goals.

Keywords: library cataloging, entity, MARC21, RDA, BIBFRAME

A könyvtári feldolgozási munkafolyamat lényeges, a hiteles információszolgáltatás szempontjából legfontosabb állomása a bibliográfiai leírások visszakereshetőségét biztosító hozzáférési pontok megállapítása. Ezek az ún. authority adatok az érvényben lévő szabályozás szerint a személy- és testületi nevek, a földrajzi nevek, valamint a címek; ezeken felül pedig különféle tárgyszavakat, tezauruszokat is használunk. Az integrált könyvtári rendszerekben a legkülönbélebb adatokat tartalmazó authority rekordok emelhetők be a leíró rekordokba, amelyek összességét authority file-nak hívják. Napjainkban az authority file már nem csupán az egyes rendszerekben, hanem az intézményeken kívül, azoktól függetlenül is létezhet – átjárást teremtve a gyűjtemények között – az ilyen szolgáltatásokat névtérnek nevezzük.

Ezek az állományok a leggyakrabban ugyancsak személynevekből (pl. VIAF, *Virtual Internet Authority File*), ULAN - *Union List of Artist Names* stb.), földrajzi nevekből (pl. GeoNames), testületi nevekből stb. épülnek, de természetesen tárgyi névterek, köznévterek is léteznek (pl. Getty). A személyek, kutatók kétséget kizáró azonosítására mindazonáltal több lehetőség is kínálkozik a VIAF mellett. Ilyen például az ISNI (*International Standard Name Identifier*), vagy az ORCID (*Open Researcher and Contributor ID*) azonosítók. A dokumentumok körében az ISBN-



azonosítás mellett új lehetőség az elektronikus formátumoknál a DOI (*Digital Object Identifier*) azonosító. Az információ- és webtechnológia fejlődésének köszönhetően napjainkban már magukat a „dolgokat” (kézzelfogható és fogalmi egyaránt) is lehetőségünk van a számítógépes feldolgozás számára egyértelműen azonosítani, ezek az URI (*Universal Resource Identifier*) azonosítók.

A könyvtári katalogizálás – amelyet ma forrásleírásként emleget a szakirodalom – elméleti alapjait az 1990-es évek végéig a Párizsi Alapelvek, gyakorlati instrukcióit pedig az európai ISBD (*International Standard Bibliographic Description*), illetve az angol-amerikai AACR2 (*Anglo-American Cataloging Rules*) szabályzatok adták. A XX. század utolsó éveiben azonban, a megsokasodó dokumentumtípusok és újfajta információhordozók megjelenése szükségessé tette, hogy a szakemberek elgondolkodjanak a bibliográfiai kontrollról és jövőjéről, amelynek egy új, általános elveket tartalmazó dokumentum megszületését köszönhetjük (ICP, *International Cataloging Principles*). Megvizsgálták azt is, hogy az információkat visszakereső használó mit vár az egyes rekordoktól, melyek azok a lényeges tematikai egységek, amelyekről tájékozódni szeretne. Ez a használóközpontú, a rekord funkcióit előtérbe helyező elemzés adta az alapját a katalogizálás új, entitásokra épülő elméletének, amelyről 1997-ben olvashattunk elsőként – ez volt az FRBR (*Functional Requirements for Bibliographic Records*), később az FRAD (*Functional Requirements for Authority Data*), majd az FRSAD (*for Subject Authority Data*). Az IFLA ezeket a különálló modelleket 2017-ben egységes szerkezetbe foglalta, megszüntetve a közöttük lévő következetlenségeket (LRM, *Library Reference Model*).

A 2010-es években a funkcionális modellekre alapozva új forrásleírási szabályzatot dolgoztak ki, amely eredetileg az AACR harmadik kiadása lett volna, végül azonban RDA (*Resource Description and Access*) néven látta meg a napvilágot. A gyakorlati megvalósításra azonban csak mostanában kezdenek hangsúlyt fektetni, mivel az informatika jelenlegi eszközei minden eddiginél hatékonyabb implementációt tesznek lehetővé. A közgyűjtemények legelterjedtebb adatcsere-formátumának, a MARC-nak (*Machine-Readable Cataloging*) ugyanis egyre több kritikával kell szembenéznie a világháló és az évről-évre egyre hatékonyabb keresőmotorok megjelenése óta. Ez a tény a már több mint fél évszázada alkalmazott technológiák, szabályok átgondolására, megújítására vagy éppen megkérdőjelezésére indította a szakembereket. Egyesek a MARC lehetőségeit igyekeznek bővíteni, míg mások a világháló előnyeit kihasználó, teljesen új lehetőségeket biztosító adatformátum implementálásán, valamint a keresőoptimalizálás általános elveinek való minél hatékonyabb megfelelésen gondolkodnak.

Az adatok összekapcsolódásán alapuló szemantikusweb-technológia fő építőkövei közé tartoznak a fent említetteken kívül az egyes tudományterületek összefüggéseinek, relációinak formalizált leírását tartalmazó ún. ontológiák, avagy szótárak. A könyvtártudomány területén is számos ilyen ismerünk – pl. a Kongresszusi Könyvtár által fejlesztett BIBFRAME (*Bibliographic Framework*) –, és napjainkban is egyre-másra jelennek meg újak, a már meglévők pedig fejlődnek, tökéletesednek.

A szótárak segítségével, RDF-tripletek (*Resource Description Framework*) formájában kifejezett adathalmazt azután egy nyilvánosan hozzáférhető tárhelyről (triplestore) bárkinek szolgáltathatjuk, akinek szüksége van rá.

Napjainkban igen erőteljes átalakuláson mennek át azok a felületek is, amelyeken a források feldolgozása, valamint szolgáltatása történik. Úgy hisszük, senkinek nem kell bemutatni a magyarországi piacon jelenlévő integrált könyvtári rendszereket (Aleph, Amicus, Corvina, Huntéka, Szikla, Szirén stb.), vagy a díjmentesen (open-source) hozzáférhető megfelelőiket (pl. a KOHA), és az OPAC (*Online Public Access Catalogue*) fogalma sem kíván bővebb definíciót. E felületek azonban az utóbbi években természetes evolúción mentek keresztül, amelynek fő mozgatórugója a világhálóba való szervezettebb beépülés, valamint az egymástól elkülönülő helyeken létező tudás összegyűjtése, integrálása volt. A letölthető, telepíthető, ún. kliensalapú rendszereket az integrált könyvtári platformok, az OPAC-okat pedig a discovery-szolgáltatások – és az ezek működését hatékonyabbá tévő SFX (*Special Effects*) – váltják fel fokozatosan.

A könyvtár ma már nem elégedhet meg a hagyományos, nyomtatott formátumú dokumentumok szolgáltatásával: tartalmait elektronikus formában, vagy éppen online is elérhetővé kell tennie a felhasználók számára. A vonatkozó szakirodalom zavarbaejtően sok kifejezéssel illeti a XX-XXI. századi szolgáltatási formákat implementáló gyűjteményeket, amelyek értelmezése akár szerzőnként, munkaközösségenként is eltérő lehet. Annyit azonban biztosan elmondhatunk, hogy napjainkban a digitális-digitalizált gyűjtemények építésének és kezelésének kiterjedt eszköztára van – repozitóriumok, adatátviteli protokollok, publikációs lehetőségek, illetve az open access-mozgalom –, amely segítséget nyújt a források feldolgozásában, tárolásában, szolgáltatásában, illetve lehetőséget biztosít az elérhető tartalmak egy felületen történő megjelenítésére, az ún. tartalom-aggregációra.

Végezetül mindenképpen érdemes szót ejteni arról a nagy informatikai átalakulási-megújulási folyamatról, amely napjainkban is zajlik az Országos Széchényi Könyvtárban. Az OKR (*Országos Könyvtári Rendszer*) néven ismert projekt számos területen tűzött ki célokat, illetve határozott meg feladatokat, amelyek megvalósítása párhuzamosan, szakmai munkacsoportokon keresztül zajlik. A legfontosabb részterületek az alábbiak:

- új integrált, országos használatra tervezett könyvtári platform bevezetése, mely lehetővé teszi a hazai közgyűjtemények együttműködését (OKP, Országos Könyvtári Platform);
- a munkafolyamatok szabványosítása, a rendelkezésre álló normatív dokumentumok magyar nyelvre fordítása, kiemelt figyelemmel az egységes szakmai terminológiára;
- nagy teljesítményű digitalizáló műhely kialakítása, ahol állományvédelmi és szolgáltatási célú digitalizálás egyaránt történhet;
- az intézmény általános informatikai infrastruktúrájának fejlesztése;
- kiemelt tartalomfejlesztési részterület, amelynek keretén belül olyan



szolgáltatások valósulnak meg, mint az egy-egy szűkebb témát feldolgozó Tudástárak, vagy a könyvtári feldolgozó munkát és az internethasználók általános tájékozódását jelentősen támogató Magyar Nemzeti Névtér.¹

A projektben a fentiek mellett nagy figyelmet fordítanak a magyar vonatkozású webhelyek archiválására és egy internet-archívum kialakítására is.

Az alábbiakban közöljük a fentebb említett témákhoz kapcsolódó legfontosabb fogalmak bővebb, részletesebb magyarázatát.

BIBFRAME: Könyvtártudományi ontológia, avagy szótár, amely a forrásleírások készítéséhez szükséges legfontosabb relációkat, viszonyokat tartalmazza. Készítői szerint ez a szótár lehet a jövőben a MARC helyett a könyvtárak egységes adatsereformátuma.

COMPASS: A Magyar Tudományos Akadémia könyvtára által gondozott, országos szintű adatbázis, amely a magyar könyvtárakban elektronikus formában elérhető, kurrens és archív tudományos adatbázisok listáját tartalmazza.

Discovery: Az online könyvtári források (a helyi nyomtatott állomány online katalógusa, helyi és kereskedelmi adatbázisok stb.) egyetlen felületen való kereshetővé tétele egyetlen keresőmező alkalmazásával (vö. Google) A legismertebb discovery-szolgáltatások a Primo, a Summon vagy az EBSCO Discovery Service (EDS), a nyílt hozzáférésű megoldások közül pedig a VuFinddal, valamint a Blacklighttal találkozhatunk a legtöbb közgyűjteménynél.

DOI (Digital Object Identifier): Az online dokumentumok körében használt egyedi azonosító rendszer, biztosítja a dokumentumok megtalálását is.

FRBR, FRAD, FRSAD: A bibliográfiai, besorolási és tárgyi besorolási rekordok funkcionális követelményeit tárgyaló elméleti modellek. Tartalmazzák a rekordok használói szempontból lényeges egységeinek (entitások) leírását és ezen egységek közötti kapcsolatokat.

ICP (International Cataloguing Principles): A 2016-ban kiadott alapelv-nyilatkozat felváltotta és határozottan kibővítette a „Párizsi alapelvek” alkalmazási területét a csupán szöveges forrásoktól az információforrások valamennyi típusáig.²

ISNI (International Standard Name Identifier): Nemzetközi szabványos névazonosító. Személyek minden kétséget kizáró azonosítására szolgál, rendszerint négyszer négy számjegyből áll.

1 OKR-projekt [elektronikus dok.] <http://www.oszk.hu/okr-projekt> [letöltés: 2019.07.18.]

2 Nyilatkozat a Nemzetközi katalógizálási alapelvekről [elektronikus dok.] URL: http://www.oszk.hu/sites/default/files/ICP_2016_magyarul_blogra_modositott.pdf [letöltés: 2019.07.18.]

Linked data: Az adatok hálózatba szervezésének, összekapcsolásának technológiája. Nem csupán a két kapcsolódó elem, de a köztük lévő viszony leírását is elvégezhetjük a segítségével. Informatikai alapkövei az egységes forrásazonosító (URI), illetve a formalizálást lehetővé tevő forrásleíró keretrendszer (RDF).

LRM (Library Reference Model): Az IFLA által kidolgozott, egységes fogalmi modell, mely összesen 11 entitást határozott meg, amelyekről a forrásleírás során "beszélni" lehet (Res, Mű, Kifejezési forma, Megjelenési forma, Példány, Ágens, Személy, Kollektív ágens, Nomen, Hely, Időtartomány).³ Az LRM megjelenése szükségessé tette az RDA alapos átdolgozását.

MARC21: Az 1998-1999-ben közzétett MARC21 a MARC legfrissebb, ma is folyamatosan aktualizált változata, a 21-es szám a XXI. századra utal. Legutóbbi módosításai figyelembe veszik az új dokumentumtípusok sajátosságait, és már az RDA-ban definiált új adatelemek leírhatóságát, valamint a korszerűbb formátumokra történő konverziót is szolgálják.

Névtér: Besorolási (authority) adatok állománya, az intézményekben végzett forrásleíró munkától többnyire függetlenül épül.

Nyílt forráskód (open source): Olyan szoftver, mely a forráskód elérhetősége révén a tetszőleges célú futtatás, a testesztelés, továbbfejlesztés, tökéletesítés jogát biztosítja a felhasználóknak. A nyílt forráskódú szoftvereken fejlesztéseket végzők egymást segítő közösségekben tevékenykednek.

OAI-PMH (Open Archives Initiative – Protocol for Metadata Harvesting): A Dublin Core formátumra építő, metaadat-begyűjtő protokoll, melynek segítségével különféle gyűjteményeket lehet egyszerű utasítások segítségével lekérdezni, vagy köztük adatcserét megvalósítani.

OJS (Open Journal Systems): Nyílt forráskódú, ingyenesen letölthető és használható, tudományos folyóiratok online közzétételére kifejlesztett folyóirat-menedzselő és publikáló rendszer.

Open Access: A tudományos közlemények nyílt hozzáférésű publikálása, mely a felhasználó számára biztosítja az olvasás és esetlegesen a további felhasználás legális, finansziális, szervezeti vagy technikai megkötések nélküli lehetőségét.⁴

ORCID (Open Research and Contributor Identifier): Nemzetközileg elfogadott szerzőazonosító kód, mely biztosítja a kutató egyértelmű azonosítását. Segítségével visszakereshető az adott kutató tudományos munkássága. Az ISNI részhalmaza.

3 IFLA könyvtári referenciamodell: A bibliográfiai információk elméleti modellje [elektronikus dok.] https://www.ifla.org/files/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712-hu.pdf [letöltés: 2019.07.18.]

4 Open Access [elektronikus dok.] URL: nkfih.gov.hu/download.php?docID=29473 [letöltés: 2019.07.18.]



Platform: a hagyományos integrált könyvtári rendszerek funkcióit kínáló, böngésző segítségével elérhető felület. Az új környezet egyszerűbbé teszi más, külső adatforrások felhasználását pl. a forrásfeldolgozásban, így megszünteti az információk elszigeteltségét. A könyvtári platformokhoz discovery-szolgáltatások kapcsolódhatnak. Könyvtári platform pl. az Alma, az OLSuite, valamint a közösségi fejlesztésű FOLIO.

RDA (Resource Description and Access): Forrásleírás és hozzáférés. Katalógizálási szabályzat, az AACR harmadik, jelentősen átdolgozott változata. A források leírásának alapjává az entitásokat és azok kapcsolatait tette, a leírás során a feldolgozó könyvtáros az RDA fejezeteit követve gyakorlatilag entitásokat ír le egymás után.

RDF (Resource Description Framework): Forrásleíró keretrendszer, amely lehetővé teszi, hogy a számítógép számára értelmezhető módon (formalizáltan) írjunk le két elemet, valamint a köztük lévő kapcsolatot. Az RDF így tehát három elemből álló "mondatokat", állításokat kezel, amelyeket tripletnek hívunk.

Repozitórium: Egyetemeken és kutatóintézményekben működő dokumentumszerver, amely tudományos anyagok archiválására és világszerte nyílt hozzáférhetővé tételére szolgál.

schema.org: A Google, a Microsoft, a Yahoo, valamint a Yandex által létrehozott kezdeményezés – egy újabb szótár –, amely segít a könyvtári rekordokat a Google-ba juttatni.

SFX: A kontextusérzékeny referenzlinkelés platformfüggetlen eszköze, amely könyvtári forrásokhoz dinamikusan generált linkhalmazokat hoz létre, s teszi mindezt az elektronikus dokumentum metaadataira támaszkodva.⁵

SPARQL: Az RDF-adatbázisok lekérdezőnyelve, adat-átalakításra is alkalmas (pl. egyik szótárról a másikra, pl. BF-schema.org)

Szemantikus web: Más néven web 3.0. A napjainkban ismert világháló egy olyan új rétege, amelyben az egyes elemek, adatok között a számítógépes feldolgozás számára is jelentéssel, értelemmel bíró, azaz informatikai eszközökkel tipizált, minősített kapcsolatok állnak fenn.

Triplestore: Adatbázis-kezelő rendszer, amely a relációs adatbázisokkal ellentétben az adatokat nem adattáblákban, hanem tripletekben tárolja (gráf-adatbázis). Ezen adatbázis lekérdező felületét (ún. SPARQL-endpoint) nyilvánosan hozzáférhetővé lehet tenni.

5 BÁNHÉGYI Zsolt: Cummings, Joel-Johnson, Ryan: Az SFX használata és használhatósága = Tudományos és Műszaki Tájékoztatás, 51. évf. (2004) 1. sz., p. 42-44.

URI (Universal Resource Identifier): Univerzális azonosító. A világon mindennek adható, legyen az kézzelfoghatóan vagy fogalmi szinten létező dolog. Ha a dolgokat és viszonyaikat egyaránt URI-kkal azonosítjuk, a számítógépes feldolgozás számára nem csupán a létezőket, de a köztük lévő kapcsolatot is formalizálni tudjuk.

VIAF (Virtual Internet Authority File): Virtuális, internetes besorolási állomány, több könyvtár (leginkább nemzeti könyvtárak) besorolási (authority) állományaiból építették. Alkalmas a személyek egyértelmű azonosítására.

Webarchiválás: weboldalak vagy webhelyek valamilyen rendszeres vagy rendszertelen időközönként megismételt mentései. A webarchívumok keresőrobotjai a linkeket követve keresik fel és mentik a weboldalakat, valamint egyéb fájlokat, amelyeket visszakereshetően tárolnak. A legismertebb webarchívum az Internet Archive.⁶

6 DRÓTOS László-KOKAS Károly - Webarchiválás és a történeti kutatások = Digitális Bölcsészeti, 1. évf. (2018) 1. sz., p. 35-53. Elektronikusan hozzáférhető a <https://doi.org/10.31400/dh-hun.2018.1.129> címen [letöltés: 2019.07.18.]