

ILÁCSA SZABINA

A könyvtári katalogizálás megújulásának lehetőségei

Napjaink egyik nagy kihívása az információkezelés területén az egységesítés, amely segítségével megvalósítható az adatintegráció a különböző információ-előállító felek között. Rengeteg mennyiségű információ található a könyvtárak webes katalógusaiban, de ezek a rekordok jórészt egymástól és a web többi részétől elzártan léteznek a weben. Ez az elzártág azt eredményezi, hogy a könyvtári katalógusok nem tudják kihasználni mindazokat a lehetőségeket, amit az Internet és a web jelenlegi technikai fejlettségünkön nyújtani tud.

A könyvtári információ leírási szabványaink frissességüket tekintve elég féltudalas képet mutatnak. Vannak nagyon friss és modern katalogizálási szabványaink (Például az elektronikus dokumentumok leírására), de a hagyományosabb dokumentumtípusokat leíró szabványaink ennél jóval régebbiek. A koruk önmagában még nem probléma, hiszen egy szabvány lényege, hogy állandóságot biztosítson. A probléma az, hogy ezek a szabványok a cédulakatalógusos időkből származnak. Megalkotásukkor a cédulakatalógus mint formátum leghatékonyabb kihasználására törekedtek. Ez a technológiai keret mára megváltozott. Ma már nem korlátoz minket, hogy a bibliográfiai rekordnak rá kell férnie egy szabványos cédulára. A technikai lehetőségek azt is megengednék, hogy bármely adatelem kereshető legyen. Mivel ez akkoriban még elképzelhetetlennek tűnt, a cédulakatalógus kori bibliográfiai leírási szabványok az adatelemek és a hozzáférési pontok megállapításánál valamiféle „azonosítási minimumra” törekedtek. Csak azokat az adatelemeket írták le, amit feltétlenül szükségesnek tartottak az adott dokumentum egyértelmű azonosításához. Ez nem feltétlenül azonos az érdeklődésre számot tartó adatelemek halmazával.

A probléma, amire oda kell figyelni az a felhasználó és az ő információs igénye és keresési módszere. Régebben a katalógusra egyfajta leltárkönyvként tekintettek, nem pedig mint a tájékoztató munka fontos eszközére.

Az MSZ 3423/1-83 definíciója szerint a leíró katalógus valamely könyvtár állományában lévő (valamely bibliográfiában regisztrált) dokumentumok betűrendbe sorolt bibliográfiai tételeinek összessége.

A leíró katalógus alapján megállapítható, hogy egy bizonyos dokumentum, amely

- szerzője és címe alapján
- ha a szerzőt a dokumentum nem nevezi meg, csak címe alapján, vagy
- ha a szerző és a cím nem alkalmas, vagy nem elegendő a dokumentum azonosításához, valamely más, a címet helyettesíthető adat alapján ismerhető fel, megvan-e a könyvtárban.

Ami nyilvánvaló ebből az idézetből, hogy az MSZ a dokumentumot helyezi a középpontba és arról szolgáltat annak azonosításához szükséges információt. Az is kiolvasható belőle, hogy az általa feltételezett katalógushasználó pontosan tudja,

hogyan melyik dokumentum(ok) kell(enek) neki és csak arra kíváncsi, hogy ez(ek) megvan(nak)-e a könyvtár állományában.

A böngészést – mint információkereső taktikát – nem támogatja, igaz akkoriban még nem ez volt az elsődleges stratégia. A webes keresőmotorok alapvetően átforgalmazták az emberek információkeresési szokásait. Hozzászoktatták őket, hogy a weben hiperlinkeken keresztül navigálnak a tartalmak között és ezt szeretnék megvalósítani a katalógusrekordok között is. Ehhez viszont az kell könyvtári részről, hogy a katalógus rekordjainkat minél több szempont alapján kapcsoljuk össze. A felhasználó ahhoz is hozzászokott a webes keresés alkalmával, hogy egyszerre rengeteg tartalomban keres. Neki az információra van szüksége és nem fontos számára, hogy honnan kapja meg. Ezért fontos, hogy a könyvtári rekordok olyan szerkezetűek legyenek, hogy könnyen integrálni lehessen más nagyobb rendszerekbe, amivel lehetővé tesszük a felhasználónak a több forrásban való egyidejű keresést. Ez az egységesítés nagyban múlik a katalogizáló szabványon is. Jelenleg vannak könyvtári területen közös katalógusaink, de csak és kizárólag könyvtári résztvevőkkel. Azt a célt kellene kitűzni, hogy egy minden tartalomszolgáltató intézmény tartalmait kezelni tudó adatbázist építsünk ki, vagy legalábbis egy közös keresőfelületet a különböző adatbázisokhoz. Ehhez viszont előbb ki kell dolgozni egy olyan szabványt és egy olyan formátumot, ami képes rá, hogy minden lehetséges formában létező a legkülönfélébb helyekről származó tartalmakat/forrásokat/dokumentumokat leírjunk és megjelenítsünk a segítségével. Valamint az is fontos, hogy ezek az adatok olyan formában kerüljenek rögzítésre, hogy aztán egy másik fél könnyen újrafelhasználhassa őket.

A következőkben azt fogom áttekinteni, hogy milyen elméleti modellek és technológiai megoldások lehetnek segítségünkre egy ilyen szabvány és formátum kidolgozásában.

A régi elméleti keretek

A bibliográfiai leírás egységesítése nem új keletű elképzés, de az első valóban számottevő hatást kiváltó kezdeményezés a múlt századig váratott magára. A Párizsi Alapelvek néven ismertté vált keretrendszert az 1961-es International Conference on Cataloguing Principles konferencián elfogadott ajánlásokból alakították ki. A konferencia legfontosabb eredményeinek a következőket tekinthetjük:

- 1.) A katalogizálás alapelveinek, a leíró katalógus funkcióinak (és a funkciók érvényesülését biztosító eszközöknek) a meghatározása.
- 2.) egységesített besorolási adatok (személynevek, testületek, címek) megválasztására és közlésmódjára (címfej, rendszó) vonatkozó legfontosabb szabályok megfogalmazása.¹

Az 1969-es koppenhágai International Meeting of Cataloguing Experts konferencián határozatként fogadták el, hogy a Párizsi Alapelveken túlmutatóan nem csak a tárgyi besorolásokat kell egységesíteni, hanem a bibliográfiai leírások formáját és tartalmát is.

Ebből a határozatból fejlődött ki az ISBD (= International Standard Bibliographic Description azaz Nemzetközi szabályos bibliográfiai leírás) program. Erre a programra azért volt szükség, mert ugyan a Párizsi Alapelvek megindított

¹ Horváth Tibor, Papp István (szerk.): Könyvtárosok kézikönyve. 2. köt. Feltárás és visszakeresés. Budapest, Osiris. 2001. 234.p.

egyfajta egységesítést, de szimplán a katalógus funkcióinak meghatározása és a besorolási adatok egységesítése nem eredményez intézmények között kicserélhető, nyelvi korlátok ellenére értelmezhető bibliográfiai rekordokat, amiket aztán könnyen lehet géppel olvasható formátumba konvertálni úgy, hogy megmaradjanak kompatibilisnek. Az ISDB a bibliográfiai leírások tartamára és formájára dolgoztak ki szabályokat. A program olyan sikeres volt, hogy az ISBD-k a nemzeti bibliográfiai központok általános szabályzataivá váltak világszerte.²

A Párizsi Alapelvek és az ISBD megfelelően erős alapokat nyújtott ahhoz, hogy az 1973-as grenoble-i IFLA konferencián meghirdetett UBC (Universal Bibliographic Control – Egyetemes bibliográfiai számbavétel) program – bár kitűzött végcél, a világbibliográfiát nem sikerült megvalósítani – ezidáig a legsikeresebb a világ dokumentumtermését regisztrálni kívánó kezdeményezések közül. A UBC a világbibliográfiát a nemzeti bibliográfiákon keresztül tartotta megvalósíthatónak. Ugyanis, ha az egyes országok a saját dokumentumtermésüket a teljesség igényével feldolgozzák és ezek rekordjait a nemzetközi szabványoknak megfelelően készítik el és teszik közzé, akkor a végeredményként a világ teljes dokumentumtermése rögzítve lesz.³

Az új elméleti keretek kialakulása

A Párizsi Alapelvek a saját korának – a cédulakatalógusok időszakának – a terméke. Így szükségszerűen nem tér ki olyan fontos dolgokra, amelyek a jelenlegi OPAC-os korunkban a katalógizálás egységesítésének fő akadályait képezik. Az IFLA ezért indítványozta a Párizsi Alapelvek felülvizsgálatát, hogy a mai digitális környezetre tartalmazzon előírásokat. A konferenciát 2003-ban Frankfurtban tartották. Ezek az új előírások nem leváltják az Alapelveket, csupán kiegészítik azokat. Az ajánlásokban szólnak mindenféle dokumentumtípusról, a tétel megválasztásról, tételformáról stb.

A tervezet alapelvei a következőket foglalják magukba:

- 1.) Alkalmazási terület
- 2.) Entitások, ismérvek és kapcsolatok
- 3.) A katalógus funkciói
- 4.) Bibliográfiai leírás
- 5.) Hozzáférési pontok
- 6.) Besorolási rekordok
- 7.) A keresési lehetőségek alapjai

Ezek az alapelvek egyrészt a régi bevált katalógizálási rendszerekre (Cutter), másrészt az IFLA által kidolgozott koncepcionális modellekre épülnek (FRBR, FRAD).

Az itt lefektetett alapelveket azzal a szándékkal fogalmazták meg, hogy útmutatóul szolgáljanak katalógizálási szabályzatok kidolgozásához. Bibliográfiai és besorolási rekordokra, valamint jelenlegi könyvtári katalógusokra vonatkoznak. Az alap-

²John Byrum: Az ISBD-k születése és újjászületése: a nemzetközi szabványos bibliográfiai leírás szabályzatainak kidolgozására és revíziójára kialakított munkafolyamatok és eljárások. In: Tudományos és Műszaki Tájékoztatás. 2001. 11.sz. 458-459 p.

³Dippold Péter: A nemzeti bibliográfiák gyűjtőköre, avagy elérhető-e a teljesség?. In: Könyvtári Figyelő. 2006. 2.sz. 261-268 p.

elvek könyvtárak, levéltárak, múzeumok és más közösségek által létrehozott bibliográfiákra és adatfájlokra egyaránt alkalmazhatók.⁴

Az FRBR modell

Az IFLA 1998-ban jelentette meg a Bibliográfiai rekordok funkcionális követelményei (Functional Requirement of Bibliographic records) című tanulmányát. A tanulmánynak két alapvető célja volt. Az első, hogy világosan meghatározott, strukturált szerkezetbe foglalja a bibliográfiai rekordokban leírt adatokat, a rekordokat használók igényeinek megfelelően. A második cél, ajánlás kidolgozása a nemzeti bibliográfiai ügynökségek által készített rekordok alapszintű funkcionalitására.⁵

A modell felvázolásakor a felhasználó által a keresési folyamat során megoldandó feladatokból indultak ki. Bibliográfiai adatok keresésekor a felhasználó feladatainak a find-identify-select-obtain (megtalálni-azonosítani-kiválasztani-megszerezni) négyest határozták meg.

- megtalálni a használó keresőkérdésében megfogalmazott kritériumoknak megfelelő dokumentumokat a használt adatok alapján (pl. a keresés során valamennyi dokumentumot adott tárgykörben; egy konkrét címen kiadott és katalogizált dokumentumot);
- azonosítani a megtalált entitást a kapott adatok alapján (pl. meggyőződni arról, hogy a rekordban leírt dokumentum megegyezik a használó által keresettel, illetve megkülönböztetni egymástól az azonos címen leírtakat);
- kiválasztani a használó kívánságának megfelelő entitást a használt adatok alapján (pl. kiválasztani a használó által ismert nyelvnek megfelelő szöveget, vagy az olyan számítógépes programot, amely megfelel a használó hardver-és operációs rendszerkövetelményeinek);
- hozzájutni a leírt entitáshoz vagy megszerezni azt a használt adatok alapján (pl. a kiadvány megrendelése, a könyvtár állományában lévő könyv kölcsönzése vagy távoli számítógépben tárolt elektronikus dokumentumhoz való online hozzáférés).⁶

A használói feladatokból kiindulva határozták meg az egyes dokumentumok különböző vetületeit, amik a használót érdekelhetik. Ez alapján entitásokat alakítottak ki, melyeket aztán csoportokba soroltak.

Ahogy az az 1. ábrán is látható, az első entitáscsoportba (szellemi vagy művészi alkotás entitásai) tartoznak a mű, kifejezési forma, megjelenési forma és a példány. A második entitáscsoportba tartozik a személy és a testület. A harmadik entitáscsoportba pedig a fogalom, tárgy, esemény, hely négyes tartozik. A modellben minden entitáshoz kapcsolódnak attribútumok is, ezek a tulajdonképpeni bibliográfiai leírás során leírt adatelemek.

⁴ Nyilatkozat a nemzetközi katalogizálási alapelvekről - Az IFLA szakértők 1. Tanácskozása által elfogadott tervezet a Nemzetközi Katalogizálási Szabályzatról 2003 Frankfurt, Németország - A „Statement of International Cataloguing Principles” magyar fordítása <http://ki.oszk.hu/kf/kfarchiv/2004/3/nyilatkozat.html> (2013-06-06)

⁵ A bibliográfiai tételek funkcionális követelményei <http://www.ifla.org/files/assets/cataloguing/frbr/frbr-hu.pdf> (2013-06-06)

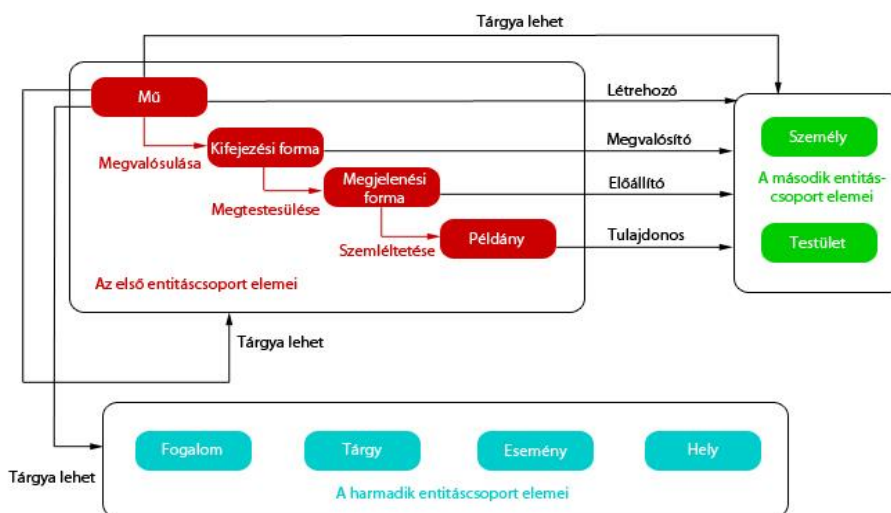
⁶ A bibliográfiai tételek funkcionális követelményei <http://www.ifla.org/files/assets/cataloguing/frbr/frbr-hu.pdf> (2013-06-06)

Az első entitáscsoport első entitása a mű, ami nem más, mint egy önálló szellemi vagy művészi alkotás. A mű egy elvont fogalomként jelenik meg a modellben, nem lehet hozzá egy a valóságban létező anyagi tárgyat kapcsolni. A szövegvariáns, fordítás nem tekinthető külön műnek, csak ugyanazon mű változatának. Új műről akkor beszélhetünk, ha az eredeti művön történt változás jelentős mértékű önálló erőbefektetést kívánt meg. Például az átiratok, adaptációk és a paródiák már önálló műveknek tekintendők. A műnek a következő ismérvei lehetnek:

- a mű címe,
- típusa (pl.: regény, térkép, szonáta stb),
- keletkezésének ideje,
- a mű környezete, összefüggése (történelmi, társadalmi, szellemi, művészi vagy más környezet, illetve összefüggés, amelyben a mű létrejött),
- célközönsége. (pl.: korosztály, képzettségi szint stb.).

Az első csoport második entitása a kifejezési forma. A kifejezési forma az a sajátos megjelenési vagy művészi forma, amelyet a mű minden egyes alkalommal felvesz, amikor megvalósításra kerül. Fordítás, hangoskönyv-változat mind egyazon mű más-más kifejezési formája. A kifejezési forma ismérvei közül néhány:

- a kifejezési forma címe,
- nyelve,
- változtathatósága (a kifejezési forma szellemi vagy művészi tartalmát várhatóan felül fogják vizsgálni (pl. valamely ideiglenes jelentés vagy tervezet, valamely címjegyzék, amely rendszeres felújításra vár.)



1. ábra FRBR entitáscsoportok

Az entitáscsoport harmadik entitása a megjelenési forma. A megjelenési forma a kifejezési forma valamilyen fizikai megtestesülése. Ez az entitás fedile a különböző kiadások és dokumentumtípusokat. Új megjelenési forma akkor keletkezik, ha az előállítási folyamat a mű fizikai formájában változást idéz elő. A megjelenési formát ismérvei közül néhány:

- kiadó,
- fizikai hordozó (pl.: papír, fa, műanyag, fém),

- megjelenési forma azonosítója (pl.: ISBN).

Az entitáscsoport negyedik eleme a példány. A példány a megjelenési forma egy darabja. Egy valóságos kézzel fogható tárgy. A példányt egyértelműen azonosító ismérvek a következők:

- példányazonosító,
- jelölések/ajánlások (dedikáció),
- a példány állapota,
- a példány hozzáférési korlátai.

A második csoportban lévő entitások (személy, testület) az első csoport entitásainak szellemi vagy művészi tartalmáért, fizikai megvalósításáért, terjesztéséért vagy megőrzéséért felelősek.

A személy entitás ismérvei közé tartoznak például:

- a személy neve,
- a személyre jellemző dátumok,
- a személynek adományozott címek,
- a személyhez kötődő egyéb megnevezések.

A testület ismérvei közé tartoznak:

- a testület neve,
- a testülethez kapcsolódó szám,
- a testülethez kapcsolt hely,
- a testülethez kapcsolt dátum,
- a testülethez kapcsolt egyéb megjelölés.

A harmadik entitáscsoportba azok az entitások tartoznak, melyek egy mű tárgyai lehetnek. A csoport elemei: fogalom, tárgy, esemény, hely. Fogalom alatt elvont fogalmakat értünk, tárgyak alatt élő vagy élettelen anyagi dolgokat, az események alatt cselekményeket és eseteket (történelmi események, korszakok stb), a helyek pedig valóságos, vagy képzeletbeli helyszínek. A harmadik entitáscsoport tagjaihoz csak egy-egy ismérv tartozik. Ezek a fogalom megnevezése, tárgy megnevezése, esemény megnevezése és a hely megnevezése.

Az első entitáscsoporton belül az entitások között hierarchikus kapcsolat áll fenn. Valamely mű egy vagy több kifejezési formában valósulhat meg. Egy kifejezési forma egy és csakis egy műnek a megnyilvánulása lehet. Valamely kifejezési forma megtestesülhet egy vagy több megjelenési formában. Egy megjelenési forma megtestesíthet egy vagy több kifejezési formát. Egy megjelenési formát egy vagy több példány szemléltethet, egy példány azonban egy és csakis egy megjelenési formát.

Az első és a második entitáscsoport tagjai között „felelősség-kapcsolatok” vázolhatók fel. Egy művet létrehozhat egy vagy több személy, illetve egy vagy több testület. Egy személy vagy egy testület viszont egy vagy több művet is létrehozhat. Ehhez hasonlóan alakul a kapcsolat a csoport többi tagja és a személy illetve testület entitás között, annyi különbséggel, hogy adott esetben megvalósítói, előállítói vagy tulajdonosi kapcsolat áll fenn.

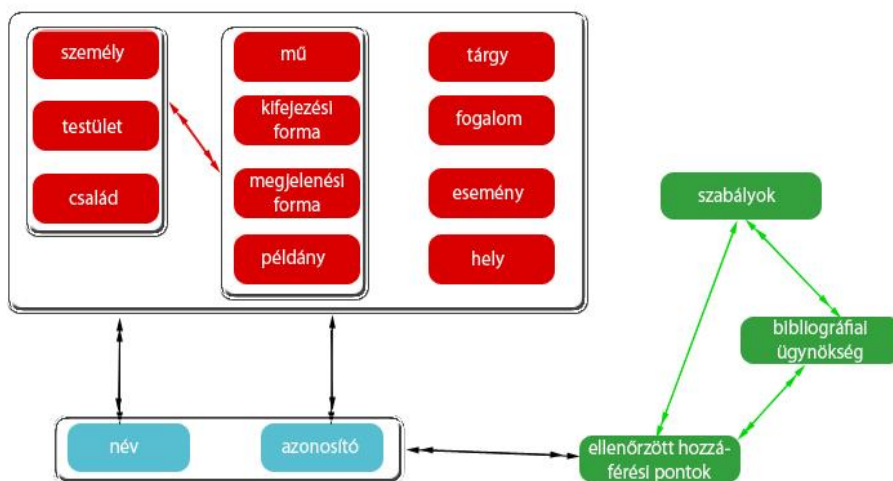
A harmadik entitáscsoport és a mű közötti kapcsolatok a következőképpen alakulnak. Valamely műnek tárgya egy vagy több fogalom, tárgy, esemény, vagy hely lehet. Ugyanúgy valamely fogalom, tárgy, esemény vagy hely egy vagy több mű tárgya is lehet.

A FRAD modell

A FRAD azaz Functional Requirement for Authority Data szintén az IFLA koncepcionális modellje. A FRAD az FRBR-t egészíti ki, mivel az FRBR nem foglalkozik a hozzáférési pontok kérdésével. Változás az FRBR-hez képest, hogy a felhasználói feladatok megváltoztak, hiszen itt most nem a bibliográfiai rekordok közötti navigálás szempontjából nézzük a használói feladatokat, hanem a hozzáférési pontok közötti eligazodás szemszögéből. A megváltozott használói feladatok a következők:

- Megtalálni, azaz információt találni az entitásról és az entáshoz kapcsolódó forrásról.
- Azonosítani, azaz megerősíteni, hogy a használó által leírt entitás megfelel a keresett entitásnak, vagy elkülöníteni két vagy több hasonló nevű entitást.
- Tisztázni, azaz tisztázni a különbséget két vagy több hasonló entitás között, vagy tisztázni a kapcsolatot a leírt entitás és a név között, amin az adott entitás ismert.
- Megérteni, azaz megérteni, hogy miért egy adott név-cím vagy név-, vagy címforma került kiválasztásra az entitás preferált nevének vagy címének.

A FRAD szerkezetét tekintve hasonló az FRBR modellhez. Szerepel benne mind a 10 FRBR entitás, de most nem mint a bibliográfiai leírás entitásai, hanem mint hozzáférési pontok. A modellben szintén 3 entitáscsoportot különíthetünk el. Ahogy az a 2. ábrán is látszik, az első entitáscsoportot a 10 FRBR entitás alkotja kiegészítve a család entitással. A modell fókuszában maga a mű áll, az összes többi entitás a művel való kapcsolatában érdekes számunkra, más szavakkal ezek az entítások a mű tárgyi hozzáférési pontjai.



2. ábra FRAD entitáscsoportok

Ami az entítások ismérveit illeti, meghagyták az FRBR-ben felvázolt entítások ismérveit és kiegészítették őket. Például a személy entitás ismérveinél megtartották a személy neve, a személyre jellemző dátumok, a személynek adományozott címek illetve a személyhez kötődő egyéb megnevezések. Kiegészítésként pedig hozzáadták a személy neme, születés helye, elhalálozás helye, ország, lakhely, munkahely, szer-

vezeti tagság, cím, alkotótevékenység során használt nyelv, foglalkozás, életrajz ismérveket.

A második entitáscsoportot a név és az azonosító alkotják, amin az első entitáscsoport tagjai ismertek. Ez a két entitás is jelen van az FRBR modellben, de míg ott attribútumokként jelentek meg, itt entitásoknak tekintjük őket. Egy entitás több néven is ismert lehet, és egy nevet több entitás is viselhet. Egy entitáshoz több azonosító is tartozhat, de egy azonosítóhoz csak egy entitás kapcsolható.

A harmadik entitáscsoportba tartoznak az egységesített hozzáférési pont, a szabályzat és a bibliográfiai ügynökség. Az egységesített hozzáférési pontok az első csoport entitásainak nevének és azonosítóinak felhasználásával készülnek. Ezeket egy bibliográfiai ügynökség által létrehozott szabályzat szerint készítik. Egy szabályzat felhasználásával több hozzáférési pont is készíthető, míg egy adott hozzáférési pontot, csak egy adott szabályzat szerint lehet elkészíteni.⁷

Amint azt láthatjuk se az FRBR se a FRAD modell nem hoz radikálisan újat. A bennük szereplő entitások és attribútumok tulajdonképpen mindegyike ismert volt már a katalogizálásban. Ami miatt mégis szükséges volt ezeket a modelleket megalakítani az az, hogy az adatelemek közti összefüggések ugyan a katalogizáló szakember számára teljesen egyértelműek voltak, de nem létezett egy olyan összefoglaló munka, amely alapján adatmodellt lehetett volna kidolgozni, így a kapcsolatok struktúrája elveszett a gépi katalogizálás/visszakeresés során. A számítógépes programoknak ugyanis adatmodellre van szükségük ahhoz, hogy kezelni tudják ezeket – az emberi intelligencia számára evidens – kapcsolatokat.

A webes keresés problémái és megoldási kísérletei: a szemantikus web

Az, hogy a számítógépes programok nem képesek kezelni az összefüggéseket, nem csak a könyvtári környezetben hátráltatja a keresést. A webes kereséskor ugyanezekkel a problémákkal találkozhatunk, sőt még halmozottabban. Míg a könyvtár egységesített adatokkal dolgozik, így annak megállapítása, hogy a találati halmazban mely elemek tartoznak a jogász Szabó Andráshoz, melyek az irodalom és egyháztörténész Szabó Andráshoz, melyek a zenész Szabó Andráshoz és így tovább viszonylag egyszerűnek mondható. A weben viszont nincs ilyesfajta egységesítés. Ha Szabó Andrásra keresünk például a Google-ben, a ránk zúdított milliós nagyságrendű találat közül akarjuk kiválasztani például mindent ami, a zenész Szabó Andrással foglalkozik már nem lesz ilyen egyszerű dolgunk. Az sem mozdít minket előre, ha a keresőkérdésünket kibővítjük Szabó András zenész-re, ugyanis a keresőmotorok működése szerint, ha az oldal nem tartalmazza a keresett szavak mindegyikét, akkor nem kerül bele a találati halmazba. Pedig könnyen előfordulhat, hogy egy releváns forrás Szabó Andrásra csak hegedűsként hivatkozik. Ilyenkor ez a forrás elveszik számunkra, mivel a keresőalgoritmus nem tudja levonni azt a következtetést, hogy az, aki hegedűs, az egyben zenész is.

A web-fejlesztőket régóta foglalkoztatja ez a probléma. A szemantikus web elképzelést először Tim Berners-Lee fogalmazta meg 2001-ben.

„A szemantikus web nem egy különálló web, csupán a már meglévő kiterjesztése, ahol az információhoz jól meghatározott jelentést társítunk, mellyel jobb ember –

⁷ Functional Requirements for Authority Data. <http://www.library.illinois.edu/cam/rda/files/FRAD.pdf> (2013-06-06)

számítógép együttműködést teszünk lehetővé. Már megtettük az első lépéseket, hogy a szemantikus webet beleszőjük a már létező web struktúrájába. A közeli jövőben ezek a fejlesztések új funkciókat tesznek majd lehetővé, ahogy a gépek egyre inkább képesek lesznek feldolgozni és „megérteni” az adatokat, amelyeket jelenleg csak megjeleníteni képesek.”⁸

A W3C szemantikus web alatt az alábbi fogalmat érti: „a szemantikus web mód-szerek és technológiák összessége, melyek segítségével elérhetővé tehetjük a gépek számára az interneten található információk jelentését, szemantikáját.”

Ahhoz, hogy a fenti célok elérhetővé váljanak, a legfontosabb dolog, hogy képesek legyünk definiálni és leírni a webes erőforrások/adatok közötti kapcsolatot. Ez kicsit más, mint a hiperlinkek jelenlegi használata, amikkel egy oldalt kapcsolunk egy másikhoz, a hiperlinkek egy kapcsolatot definiálnak a jelenlegi weboldal és a céloldal között. Egy nagyon fontos különbség a hiperlink és a szemantikus web között, hogy az utóbbinál ilyesféle kapcsolatot lehet definiálni két bármilyen forrás között, nincs értelme a jelenlegi weboldal fogalmának ebben a környezetben. Egy másik fontos különbség, hogy a definiált kapcsolat (maga a link) is el van nevezve, míg a hagyományos linkeknél nem ez a helyzet így a szerepük leredukálódik az emberi használatra. A fent említett kapcsolatok definiálása lehetővé tesz számunkra egy jobb és automatikusan működő adatcserét. Az RDF (Resource Description Framework) ami a szemantikus web alapköve, ennek az adatcserének a megvalósításához ad egy formális keretet.⁹



3. ábra

A szemantikus web fontos alapköve az RDF (Resource Description Framework – forrás leíró keretrendszer). Az RDF úgynevezett tripletekből áll. A triplet részei a szubjektum, a predikátum és az objektum. Ezek megfelelnek egy hagyományos természetes nyelvi mondat alany-állítmány-tárgy felépítésének. Az RFD tripletet a 3. ábrán látható módon szokás általában szemléltetni.¹⁰

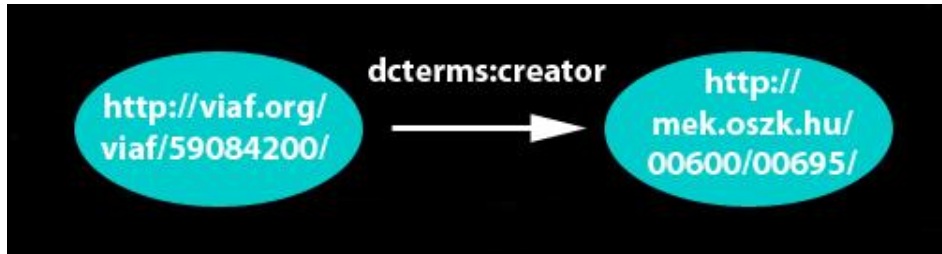
Mindegyik „mondatrész” lehet ember által olvasható természetes nyelvű karaktersorozat (literál), de ez általában csak szemléltetés céljából használatos, mivel a literálokat a programok nem értik. Így a gyakorlatban működő RDF tripleteknél a

⁸ Tim Berners-Lee - James Hendler - Ora Lassila: The Semantic Web <http://www.scientific-american.com/article.cfm?id=the-semantic-web&page=2> (2013-06-06)

⁹ W3C Semantic Web / Frequently Asked Questions / 1.3. What are the major building blocks of the Semantic Web? <http://www.w3.org/2001/sw/SW-FAQ#whatarebuildingblocks> (2013-06-06)

¹⁰ Tóth Máté: Könyvtárak a szemantikus web világában. In: Könyvtár Figyelő. 2010. 3.sz. 413-438. p.

források és a relációk is URI-k (Universal Resource Identifier = egyetemes erőforrás azonosítók). Az URI-kat elsősorban gépek értik. Az URI egy webes erőforrást azonosító karaktersorozat. A gyakorlatban 100-ból 99 esetben az URI egy sima, jól megszokott URL, a fennmaradó 1 esetben pedig nagyjából úgy néz ki, mint egy URL csak más prefixumot használ.¹¹ A 3. ábrán látható gráf valóság közelebb megjelenítése a 4. ábrán látható.



4. ábra

Az irányított gráfos ábrázolást csak a könnyebb ábrázolás és megértés miatt szokás használni az RDF esetében. Természetesen a valóságban nem így néznek ki az adatkapcsolatok. Az RDF-et – mivel platformfüggetlen adatszerenyelv – többféle szintaxisban is lehet írni. A más ismert RDF triplet XML szintaxisban az 5. ábrán látható.

```
<?xml version="1.0" ?>
<rdf:RDF xmlns:rdf="http://www.w3c.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description rdf:about="http://mek.oszk.hu/00600/00695/">
    <dc:creator rdf:resource="http://viaf.org/viaf/59084200/">
    </dc:creator>
  </rdf:Description>
</rdf:RDF>
```

5. ábra

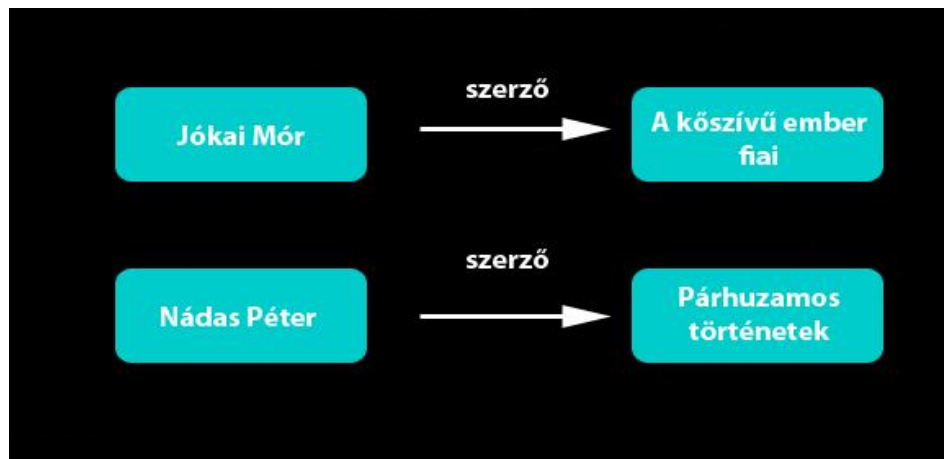
Megfelelő összefogással RDF-ben le lehetne írni minden webes forrást és megvalósítható lenne a szemantikai alapú kereshetőség. Ez még jelenleg nincs így. Bár vannak a szemantikus web terén nagyon szép eredmények, még nagyon kevesen használják a kidolgozott ajánlásokat, így nem igazán használhatók ki a benne rejlő lehetőségek.

RDF-ben megjeleníthetünk nem RDF-ben leírt forrásokat is, ha azok RDF-szerű módon/logikával vannak rögzítve. Például egy táblázat tartalma általában viszonylag könnyen és automatikusan átírható RDF-be, hiszen az adok rögzítési módja ezt lehetővé teszi. Az adatok jól el vannak különítve mezőkbe. Egy mezőben csak egyféle adat található. A táblázat oszlopainak címsora pedig információt szolgáltat arról, hogy az adott oszlop adatai milyen kapcsolatban állnak a többi oszlop adataival.

¹¹ Dorothea Salo: RDF, RDA, and other TLAs. <http://www.slideshare.net/cavlec/rdf-rda-and-other-tlas> (2013-06-06)

Mű címe	Szerző
A köszívű ember fiai	Jókai Mór
Párhuzamos történetek	Nádas Péter

1. táblázat



6. ábra

A szemantikus web szabványok jelenleg még nem terjedtek el eléggé ahhoz, hogy igazán látványos változást indíthassanak el. De vannak már működő alkalmazásai. Ilyen például a Tim Berners-Lee által nemrégiben készített <http://data.gov.uk/>.¹² Ez az oldal az Egyesült Királyság kormányzatáról szolgáltat könnyen újrafeldolgozható adatokat. Ezeknek segítségével különböző hasznos alkalmazásokat lehet készíteni. Az oldalon több megtalálható ezekből az alkalmazásokból. Példaként hoznám a Locatable nevű alkalmazást, amely megkönnyíti a „milyen környékre költözzünk” probléma megoldását. Az oldalon megtalálható adatokra építve egy térképen keresgélve megnézhetjük az adott környék lakhatási szempontból fontosabb jellemzőit, mint például tömegközlekedés, bűnözési ráta stb. Ezen paraméterek alapján kereshetünk is, hogy megtaláljuk azt a környéket, ahol lakni szeretnénk.

A data.gov.uk nagyon jól példázza, hogy mit szeretne a szemantikus web iránt elkötelezett elérni. Ezek az adatok eddig is szabadon hozzáférhetők voltak mindenki számára az alkalmazások megírása előtt is. Viszont az adatok szemantikus web szabványoknak megfelelő rögzítése miatt lehetőség nyílt rá, hogy ezeket az adatokat automatikusan felhasználva valami újat hozzanak létre, jelentősen lecsökkentve a keresésre fordított időt és nem utolsó sorban jobb felhasználói élményt nyújtva.

Miért lenne hasznos ezeken a modelleken és technológiákon alapuló katalogizálási szabványt és formátumot kidolgozni?

Az FR metaadatcsaládon alapuló katalogizálási szabványok azért lennének hasznosak a könyvtárak számára, mivel sokkal célirányosabb keresést tennének

¹² DATA.GOV.UK – Opening up Government / About us / Who is involved in the project? <http://data.gov.uk/about-us> (2013-06-06)

lehetővé. Jelenleg (az FRBR modell szerint) megjelenési formákat viszünk fel a katalógusba. A hatékonyságvesztés abból adódik, hogy az ily módon felvitt rekordok nem feltétlenül arra a kérdésre adják meg a választ, amit a használó megfogalmazott. Például annak a kérdésnek a megválaszolása, hogy hány példányban van meg a könyvtárban a Hamlet Arany János-féle fordítása eléggé munkaigényes. A katalógus ugyan kilistázza nekünk az összes olyan megjelenési formát, ami tartalmazza az említett kifejezési formát, viszont a példányok száma csak úgy állapítható meg, ha minden rekordnál külön megnézzük, hogy hány példány van hozzá csatolva az egyes megjelenési formákhoz. Ugyanis a jelenlegi bibliográfiai leírásokban a 4 FRBR szintből csak 3 attribútumai szerepelnek. A példányhoz kapcsolódó attribútumok a katalógusban jellemzően a bibliográfiai leírástól teljesen elkülönülve jelennek meg.

A szemantikus web alapú adatrögzítés alkalmazásának előnye pedig az lenne, hogy a könyvtári rekordokat kimozdítaná a jelenlegi elszigetelt állapotukból. Az adatok így nem csak láthatóak lennének a szemantikus alapon működő webes keresők számára, de egyben újrafelhasználhatóvá is válnának a könyvtári metaadatok. A könyvtárak rendelkeznek a világ egyik legnagyobb rendszerezett, minőségi metaadatnyilvántartásával. Ennek szemantikus web kompatibilis formába történő átírása nagy mértékben hozzájárulna ahhoz, hogy a szemantikus webes szabványokon alapuló adatrögzítés elérje azt a bizonyos „kritikus tömeget”.

A szemantikus webben rejlő lehetőségeket az IFLA is felismerte. 2012 májusában a Funkcionális követelmények bibliográfiai metaadatmodell-család névtér-definícióit elérhetővé tette RDF formátumban az Open Metadata Registryben.¹³

Könyvtári részről több próbálkozás is született az FRBR (REICAT – az új olasz katalogizálási szabvány¹⁴) és a szemantikus web (a svéd nemzeti könyvtár katalógusának konverziója¹⁵) nyújtotta lehetőségek kiaknázására. De eddig csak a Joint Steering Committee for Development of RDA fejlesztésében megjelent Resource Description and Access (RDA) az, amely mind az FR modelleket, mind a szemantikus webes ajánlásokat felhasználta. Az RDA adatmodelljét az FRBR és a FRAD modell alapján dolgozták ki és a szabvány szinte teljesen át is vette a modellek fogalmi rendszerét.¹⁶ A szabványhoz saját formátumot (BIBFRAME) terveztek, ami szemantikus web alapú.¹⁷ A szabványt bevezetése mellett olyan meghatározó könyvtárak kötelezték el magukat, mint a Library of Congress, a British Library vagy a Deutsche Nationalbibliothek. Még korai lenne megjósolni, hogy az RDA mekkora körben fog elterjedni és hogy egyáltalán az RDA lesz-e az a szabvány, ami meghatározza majd a jövő katalogizálását, de azt kijelenthetjük, hogy az FR modellek és a szemantikus web összehangolása lehet az alapja a katalogizálás modernkori forra-

¹³ Dudás Anikó: Nemcsak weben lenni, hanem webből lenni – A Funkcionális követelmények (FR) metaadatmodell-család névterei és a szemantikus web. In: Könyvtári Figyelő 2013. 1. sz. 45-64p.

¹⁴ Antonella Trombone: The New Italian Cataloguing Rules (REICAT) and the UNIMARC standard: open problems and proposals for the application in library catalogues. <http://conference.ifla.org/past/ifla75/135-trombone-en.pdf> (2013-06-06)

¹⁵ Tóth Máté: Könyvtárak a szemantikus web világában. In: Könyvtár Figyelő. 2010. 3.sz. 413-438. p.

¹⁶ Dudás Anikó: Forrásleírás és hozzáférés. Az új angol-amerikai katalogizálási szabályzat (RDA) és kritikája. In: Könyvtári Figyelő, 2012. 4.sz. 740.p.

¹⁷ <http://bibframe.org/> (2013-06-06)

dalmának. Erre a keretrendszerre építve megvalósulhat a más közösségek által is hasznosítható, többszintű, strukturált bibliográfiai adatrögzítés.

Irodalomjegyzék

- John Byrum: Az ISBD-k születése és újjászületése: a nemzetközi szabványos bibliográfiai leírás szabályzatainak kidolgozására és revíziójára kialakított munkafolyamatok és eljárások. In: Tudományos és Műszaki Tájékoztató. 2001. 11.sz. 458-459 p.
- Dippold Péter: A nemzeti bibliográfiák gyűjtőköre, avagy elérhető-e a teljesség?. In: Könyvtári Figyelő. 2006. 2.sz. 261-268 p.
- Dudás Anikó: Forrásleírás és hozzáférés. Az új angol-amerikai katalogizálási szabályzat (RDA) és kritikája. In: Könyvtári Figyelő, 2012. 4.sz. 740.p.
- Dudás Anikó: Nemcsak weben lenni, hanem webből lenni – A Funkcionális követelmények (FR) metaadatmodell-család névterei és a szemantikus web. In: Könyvtári Figyelő 2013. 1. sz. 45-64p.
- Horváth Tibor, Papp István (szerk.): Könyvtárosok kézikönyve. 2. köt. Feltárás és visszakeresés. Budapest, Osiris. 2001. 234.p.
- Tóth Máté: Könyvtárak a szemantikus web világában. In: Könyvtár Figyelő. 2010. 3.sz. 413-438. p.

Elektronikus források

- A bibliográfiai tételek funkcionális követelményei <http://www.ifla.org/files/assets/cataloguing/frbr/frbr-hu.pdf> (2013-06-06)
- Antonella Trombone: The New Italian Cataloguing Rules (REICAT) and the UNIMARC standard: open problems and proposals for the application in library catalogues. <http://conference.ifla.org/past/ifla75/135-trombone-en.pdf> (2013-06-06)
- Tim Berners-Lee – James Hendler – Ora Lassila: The Semantic Web <http://www.scientificamerican.com/article.cfm?id=the-semantic-web&page=2> (2013-06-06)
- <http://bibframe.org/> (2013-06-06)
- DATA.GOV.UK – Opening up Government / About us / Who is involved in the project? <http://data.gov.uk/about-us> (2013-06-06)
- Functional Requirements for Authority Data. <http://www.library.illinois.edu/cam/rda/files/FRAD.pdf> (2013-06-06)
- Nyilatkozat a nemzetközi katalogizálási alelvekről – Az IFLA szakértők 1. Tanácskozása által elfogadott tervezet a Nemzetközi Katalogizálási Szabályzatról 2003 Frankfurt, Németország – A „Statement of International Cataloguing Principles” magyar fordítása <http://ki.oszk.hu/kf/kfarchiv/2004/3/nyilatkozat.html> (2013-06-06)
- Dorothea Salo: RDF, RDA, and other TLAs. <http://www.slideshare.net/cavlec/rdf-rda-and-other-tlas> (2013-06-06)
- W3C Semantic Web / Frequently Asked Questions / 1.3. What are the major building blocks of the Semantic Web? <http://www.w3.org/2001/sw/SW-FAQ#whatarebuildingblocks> (2013-06-06)