

Könyvtárak a szemantikus web világában

TÓTH Máté

Az alábbi tanulmány a szerző „A szemantikus web. Könyvtárak a szemantikai paradigmában.” (Bp. : ELTE Irodalomtudományi Doktori Iskola, 2010. 180 p.) című, nemrégiben sikeresen megvédett disszertációja alapján készült lapunk számára.

Bevezetés

A kutatás könyvtári szempontból vizsgálja a webes szolgáltatások új generációja által generált lehetőségeket és kihívásokat. A szemantikus web tárgykörébe tartozó fejlesztések már több mint tíz éve jelen vannak a tudományos és technológiai kutatásokban, de a könyvtáros szakma részéről nem született olyan áttekintés, amely saját szempontjából elemezte volna a jelenséget. Paradoxon, hogy míg a szakma mind erőteljesebben hangsúlyozza, hogy a világhálón lévő tartalmak rendezésére ellenőrzött szótárakra, tradicionális könyvtáros készségekkel rendelkezők tömege-

inek munkájára lenne szükség, a szemantikus web fogalma többnyire ismeretlenül cseng az átlag szakmabeli fülében. (Vö. Burke 2009; Macgregor 2008) A szemantikus technológiák ugyanakkor mind több alkalmazásban érhetőek tetten világszerte. A célkitűzésünk tehát egy könyvtári szempontú áttekintés készítése a szemantikus technológiákról, valamint a hazai és nemzetközi szemantikusweb-fejlesztésekről, majd az ezek nyomán megjelenő technológiai környezet jellemzői alapján jövőkép felrajzolása a könyvtárak számára.

Nem tekintjük feladatunknak a szemantikus technológiák könyvtári alkalmazásával kapcsolatos technikai jellegű kérdések részletes tárgyalását és megoldását.

Tézisünk, hogy a könyvtár az általa birtokolt (metaadat- és egyéb) tartalmak szemantikus technológiákkal kompatibilis közzétételével, tudásszervezési rendszerek készítésével, valamint a szemantikusweb-szabványok támogatásával járulhat hozzá egy globális hálózati metaadat-infrastruktúra megvalósításához.

A szemantikus web víziója

Mindenekelőtt ismerkedjünk meg a szemantikus web elképzelés lényegével. *Tim Berners-Lee* és kollégái 2001 májusában tették közzé a *Scientific American* hasábjain az intelligens világháló alapvetését, amelyben a következő példával illusztrálják a szemantikus web fejlesztések várható eredményét. Az elképzelt történet szerint Pete és Lucy testvérek, akik édesanyjuknak fizioterápiás kezelést keresnek kétheti rendszerességgel. Az intelligens kereső kikeresi az édesanyjuknak előírt kezelést nyújtó szolgáltatók közül azokat, amelyeknél érvényes a mama biztosítása. Kiválogatja azokat, amelyek a lakhelyének 20 mérföldes körzetében található, és az értékelésük „kiváló” vagy „nagyon jó”. Ezt követően összeegyezteti az időpontot Lucy és Pete naptárával, hogy nekik is megfelelő időben legyen a kezelés. Az intelligens kereső vázolja a tervet. Pete-nek nem tetszik, mert éppen csúc-

forgalom idején kellene autózni, ezért megismétli a keresést szigorúbb kritériumokkal. Szinte azonnal jön a válasz a módosított tervvel, de a gép figyelmezteti, hogy akkor néhány kevésbé fontos találkozóját át kell tennie más időpontra. (Berners-Lee, Handler, Lassila 2001)

Az elképzelt történet és a vázolt technológiai megoldások recepciója a szakirodalomban nagyon gazdag. *Passin* (2004) nyolc fő csapásirányt vázol fel a szemantikus web vízió első interpretációiból. Ezek a fenti történetből kiragadott elemekre alapulnak.

- ◎ *Az információ indexelése és visszakeresése.* Az információ-visszakeresésben a szemantikus web túlmegy a tárgyszók és az alfabetikus indexek mélységén, így a használók képesek fogalmak és kategóriák szerint keresni.
- ◎ *Metaadatok.* A szemantikus web metadatok tömege, amelyek a weben lévő adatok közötti hatékonyabb keresést teszik lehetővé.
- ◎ *Magyarázó jegyzetek, kommentárok.* A szemantikus web legfőbb újítása a felhasználó által generált tartalom (kommentárok, bejegyzések).
- ◎ *Egy nagy együttműködő adatbázis.* Az adatok legnagyobb részét adatbázisokon kívül tárolják a weben. A tárolt adatok leírásának és visszakeresési lehetőségeinek egységesítése nyomán a szemantikus web egy nagy virtuális adatbázisként is értelmezhető.
- ◎ *Az adatok gépi visszakeresése.* Ez a megközelítés az adatok visszakeresésének az automatizálására koncentrál. E szerint nem kell az oldalakat egyenként betölteni és azokról az adatokat kinyerni, hiszen az adatok formátuma és hozzáférési lehetőségei lehetővé teszik, hogy azokat a számítógép emberi közreműködés nélkül kezelje.
- ◎ *Szolgáltatások.* A szolgáltatások géppel olvasható adatokat tesznek közzé, így a számítógép az ember helyett végzi az egyeztetéseket.
- ◎ *Szolgáltatások felderítése.* Azon szolgáltatások felderítése, amelyek lehetővé teszik a szabványok szerinti hozzáférést.
- ◎ *Intelligens „titkár”.* Olyan intelligens számítógépes alkalmazás, amely az ember helyett

cselekszik gyakorlati ügyek intézése során (például repülőjegyet vagy szállodai szobát rendel). A szemantikus weben ebben a megközelítésben emberek intelligens „titkárainak” interakciója zajlik. (Passin 2004)

Minden megközelítés hordoz igazságot. Összefoglalva azonban a szemantikus web célja egy olyan globális hálózati metaadat-infrastruktúra létrehozása, amely lehetővé teszi a világhálón lévő adatok integrálását, a közöttük lévő kapcsolatok definiálását és jellemzését, illetve az adatok értelmezését.

Berners-Lee és kollégái (2001) az intelligens világháló felépítését olyan lépcsős modellben képzelik el, amelynek a legalsó szintjén az egyedi adatok megragadására, következő szintjén a közös szintaxisra, harmadik szintjén a szemantikus tartalomra, majd a logikára vonatkozó metaadatszabványok foglalnak helyet. Az egyes rétegekben nyelvek találhatók, amelyeknek az adatok egyértelmű címzésétől egészen a tudás bonyolult struktúráinak modellezéséig számos feladatot kell ellátniuk.

A távlati elképzelések alapján az intelligens háló az önleíró dokumentumra épülve a metaadatok segítségével a formális logikai szabályok szerinti következtetések levonására lesz képes.

Szakirodalmi áttekintés

Az intelligens világháló nagyon intenzíven hatott a hazai és nemzetközi könyvtártudományi kutatásokra, amelyek főbb eredményeit és csapásirányait az alábbiakban foglalom össze. Az áttekintésben csak megemlítem a szemantikus webes technológiákról szóló általános monográfiákat. Magyar nyelven *Gottdank Tibor* (2005), illetve *Szeredi Péter – Lukácsy Gergely – Benkő Tamás* (2005) írtak e témában könyvet. Angol nyelven a teljesség igénye nélkül a következők az ismertebb művek szerzői: *Michael C. Daconta – Leo J. Obrst – Kevin T. Smith* (2003), *Thomas B. Passin* (2004), *Liyang Yu* (2007), *H. Peter Alesso – Craig F. Smith* (2009).

A Tudományos és Műszaki Tájékoztatás c. folyó-

irat több cikket is közölt a szemantikus webről, azonban egyik sem kifejezetten a könyvtári alkalmazási lehetőségeket mutatja be. *Fülöp Csaba, Kovács László és Micsik András* (2004) ismertetik a szemantikus web elképzelést, valamint az európai *Cores* projektet, amely azt a célt tűzte ki, hogy bővítse a metaadatszabványok és megoldások együttműködését a metaadatsémák nyilvántartásával. A szemantikus világháló felé vezető úton az ellenőrzött szótárak, a szabványos leíró nyelvek és kiemelten a taxonómiák szerepére hívja fel a figyelmet *Horváth Zoltánné* (2006). A könyvtárosokat leginkább a tudás osztályozási rendszerekben való formalizálása érdekli. A szemantikus web fejlesztésében a könyvtárosokra elsősorban a tudásszervezési rendszerek építésének feladata hárul. A cikk bemutatja a szemantikus web alapötletét, kitekintéssel a tématerkép-technológiára is. *Ungváry Rudolf* (2004) a szemantikus web legfőbb alkotóelemét, az ontológiákat, azok tezauruszokhoz képest megfogalmazható különbségeit, valamint a bennük leírt generikus hierarchialáncokat mutatja be könyvtáros szempontból. A tanulmány rávilágít arra, hogy a könyvtártudomány a szemantikusweb-fejlesztések kulcs-területe. „*Fontos, hogy kapcsolatot teremtsünk a hagyományos és az új szóhasználat között, és megvilágítsuk, hogy a könyvtári világ az információgazdaság legkorábbi intézménye, melynek számos eredménye felhasználható az ismeret-technológiában, s egyúttal sikerrel integrálható a tartalomiparban (ahogy a webes információszolgáltatók saját működési környezetüket nevezik).*” (Ez a felismerés egyébként a szemantikus web vízió publikálását követően szinte azonnal megjelent a nemzetközi szakirodalomban is; ld. Goossens 2003.)

A szemantikus web könyvtári trendjeinek elemzésére tettem kísérletet a 2006. évi AGRIA MÉDIA konferencián. A digitális könyvtár definíciói fényében a cikk azt a következtetést fogalmazza meg, miszerint a szemantikusweb-technológiák fejlesztése előrevetíti a könyvtár-automatizálás időszakában elindult tendenciák erősödését az információ legelemibb hordozó egységeitől

kiindulva a rendszer teljes komplexitásáig. Rövid távon a dokumentumok feldolgozása során született metaadatok közötti szemantikus átjárhatóság biztosítása jelent kihívást, hosszú távon azonban egy egészen új szolgáltatási paradigmát kell felrajzolnunk. (Tóth 2007)

Szakadát István (2007b) azt a kérdést teszi fel, hogy lesznek-e, illetve mire lesznek jók a könyvtárosok a jövőben. A Web 2.0-s jelenségek fölöslegessé tehetik a könyvtárosok munkáját, míg a hiányzó szemantikai információk pótlására könyvtárosok, professzionális archivátorok munkájára, valamint gépi szemantikára, gépi tudásrendszerbe való beépítésére van szükség. Egyelőre nem látni, hogy a szemantikus web megvalósítására lesz-e bármikor esély, terjednek viszont a folkszonómikus rendszerek, amelyek nélkülözik a szemantika megragadását. A könyvtárosok szerepe a jövőben az automatikus szemantikai feltáráshoz a gépek „megtanítása” arra, hogy miként kell feldolgozni a szövegeket.

A Könyvtörténeti és Könyvtártudományi Szakkönyvtár katalógusa szerint eddig mindössze egyetlen könyvtártudományi-informatikai szakdolgozat született Magyarországon a szemantikus webről. Ezt *Vas Beatrix* (2005) készítette a Szegedi Tudományegyetem Könyvtártudományi tanszékén. A dolgozat áttekintést ad a hagyományos tartalomfeltárás korlátairól, és arról, hogy elektronikus környezetben milyen alternatívái vannak az ismeretek feldolgozásának. Ismerteti a szemantikus web legfontosabb építőelemeit, az RDF-et és az OWL-t. Bár a szerző az új technológiák bemutatását vitaindító céllal tette, az írásnak sajnos, nem lett visszhangja.

A külföldi szakirodalomban a szemantikus web a kezdetektől az érdeklődés homlokterében volt. (Jellemző azonban, hogy a Web 2.0-nak, illetve a fogalom nyomán kibontakozó Könyvtár 2.0 jelenségnek sokkal terjedelmesebb irodalma van. A könyvtári világ számára talán világosabb üzenet volt a közösségi tartalomépítés, mint egy elvont metaadat-birodalom építésének a szükségessége.)

Az Information Research közölte – Tim Berners-Lee és kollégái (2001) cikkét követően

– az első áttekintést, amely a könyvtártudományi kutatás jövőbeli irányait igyekszik vázolni az intelligens világhálóról szóló elképzelés fényében. Lu, Dong és Fotouchi (2002) négy lehetőséget soroltak fel:

1. webes szolgáltatások fejlesztése;
2. ágens-alapú megosztott programozási paradigma megjelenése;
3. szemantikus keresőmotorok fejlesztése;
4. szemantikus digitális könyvtárak építése.

A lehetőségek mellett kihívásként vették számba:

1. az ontológiák fejlesztését,
2. a formális logikák beépítését a szemantikus web nyelveibe, valamint,
3. a bizalomépítést.

A 2002-ben vázolt lehetőségek és kihívások – kevés módosítással – mind a mai napig helytállóak.

A Library Management 2005-ben különszámot adott ki a témában. A kötet motivációját az a világszerte zajló intenzív kutatási tevékenység jelentette, amely a kollektív emberi tudás megragadását és kifejezését tűzte ki célul. A különszám célként jelölte meg, hogy a digitális könyvtárak és a szemantikus web kutatása számára iránymutatásul szolgáljon. Általánosítva a következő feladatok megoldásában lehet jelentősége az új generációs világhálóknak:

1. szemantikus interoperabilitás;
2. rutinfeladatok elvégeztetése ágensekkel;
3. felhasználói felületeken való böngészés, keresés, szűrés lehetőségeinek javítása.

A különszám cikkei egyrészt bevezető írások, amelyek segítik az olvasókat abban, hogy megértsék a szemantikusweb-fejlesztéseket, valamint azok digitális könyvtárakra vonatkozó kihívásait, másrészt olyan kutatások, amelyek digitális könyvtári környezetben alkalmazzák az intelligens világháló technológiáit. (Lytras, Sicilia, Davies, Kashyap 2005) A tematikus szám összefoglalja a digitális könyvtárak szempontjából releváns szemantikusweb-technológiákat. (Sure, Studer 2005) A könyvtárak önmagukat tudásmenedzsment központként definiálva a kis- és kö-

zepes vállalkozások számára a szemantikusweb-technológiák használatával képesek olyan szolgáltatásokat nyújtani, amelyekkel azok versenyképesebbé tehetők. Mivel e webtechnológiák többnyire könyvtáros szakértelmet igényelnek, a digitális könyvtárak megőrizhetik jelentőségüket a 21. században is. Ennek egyik modellje az LKMC (Libraries as Knowledge Management Centers; Parker, Nitse, Flowers 2005). Warren és Alsmeyer (2005) négy problémát fogalmazott meg a digitális könyvtári területen, majd azt kutatták, hogy a szemantikus technológiák miként lennének képesek ezeket kezelni. Ezek:

1. a teljes interoperabilitás megvalósítása,
2. az objektumok és a repozitóriumok leírása,
3. a multimédia gyűjtemények kezelése, valamint
4. a felhasználói felületek és az ember-gép interakció fejlesztése.

A *Cataloging & Classification Quarterly* 2007-ben szentelt különszámot a témának. A tematikus szám tanulmányai párhuzamosan megjelentek a *Knitting the semantic web* c. könyvben. A kötet 13 tanulmányt tartalmaz. Az első rész írásai teoretikus megközelítést adnak, és azokat a technológiákat tárgyalják, amelyek relevánsak a könyvtárosok szempontjából, a második különböző szakterületeken készített konkrét alkalmazásokat tekint át. (Greenberg, Méndez 2007) *Jane Greenberg* (2007) az alapvető könyvtári funkciókat alkalmazza a szemantikus web kontextusában a következők szerint. Az állománygyarapítás a szemantikus web tartalmak kiválasztása. A katalogizálás tartalmak szemantikus reprezentálása. A tájékoztatás szemantikusweb-szolgáltatás. A dokumentumforgalmazás szemantikusweb-erőforrások használata. A cikk szerint a megfeleltethetőség ellenére szakadék van a digitális könyvtári és a szemantikusweb-kutatások között, amelyet több okból is fel kellene számolni. A szemantikus webes közösség rendszerint nem rendelkezik azokkal a készségekkel, amelyekkel a könyvtárosok igen, és amelyek a tudásreprezentáció során szükségesek lennének. A könyvtárosok lassú-

nak bizonyultak a szemantikusweb-technológiák adaptálásában. Az intelligens világháló egyelőre nagyon kevés konkrétumot hordoz, így a könyvtárosok nem is tudják elképzelni, hogyan tudnának bekapcsolódni ebbe a munkába. Megoldás lehet a könyvtárosok intenzívebb részvétele a W3C (World Wide Web3 Consortium) tevékenységében, valamint közös rendezvényeken, konferenciákon való tapasztalatcsere.

Szintén a tradicionális könyvtári világ és a szemantikus web közötti párhuzamokat kutatja *McCathie* és *Méndez* (2007) a katalogizálás és osztályozás szempontjából. Javaslatot tesznek a szemantikusweb-szabványok (RDF alapú) katalogizálásra. Az RDF nagyon egyszerű szintaxis, amely egyszerre képes a hagyományos katalógusok kifejezőerejét megtartani, és támogatni a digitális könyvtárak metaadatainak cseréjét.

Karen Coyle (2008) azt elemzi, hogy miként használhatnák ki a könyvtárak a szemantikus web lehetőségeit. A fejlesztések kulcsterülete a jelentés géppel olvasható formában való kódolása. A könyvtárak ezért a gyakorlati fejlesztés szempontjából kulcspozícióban vannak, hiszen a katalógusaikban tömegesen áll rendelkezésre szemantikailag kódolt információ. A könyvtárosoknak két teendőjük van a jövőben. Egyrészt a rendelkezésre álló információ tömeget át kell kódolni a szemantikusweb-szabványokhoz illeszkedő formátumba, másrészt az adatbázisokban tárolt információt meg kell nyitni, hogy azok más webes erőforrásokkal átjárhatóak legyenek. *A Library Review* 2008-ban szánt tematikus számot a szemantikusweb-kutatások könyvtári vonatkozásainak. A különszám célként tűzte ki, hogy elősegítse a digitális könyvtári és a szemantikusweb-kutatások közötti párbeszédet. A digitális könyvtárak érett információs szolgáltatások, ellentétben a szemantikus webbel, amely rengeteg ígéretet hordoz, de még mindig a fejlesztések kezdeti stádiumában van. Az intelligens világháló víziójának legnagyobb ígérete, hogy a heterogén erőforrások felkutatását és adatok közötti együttműködést fogja segíteni. A kulturális intézmények, adattárak, könyvtárak számára lehetőséget teremt a saját tartalmak

hatékonyabb „tálalására”. Mindezek ellenére a könyvtárosok nem foglalkoztak a szemantikus webbel a valós súlyának megfelelően. A változást a Web 2.0-ról szóló diskurzus megjelenése indította el. (MacGregor 2008)

A tematikus szám ezután konkrét alkalmazásokról szóló cikkeket közöl. A bevezető írás a gyakorló könyvtáros szemszögéből vizsgálja a szemantikusweb-fejlesztések következményeit, illetve megkísérli áthidalni az elméleti megközelítések és az átlag szakmabeli szempontja között mutatkozó különbségeket. A kiváló ötletek nem kerültek át a mindennapi munkába, így a gyakorló könyvtáros számára a szemantikus webnek egyelőre nincs jelentősége a napi rutin szempontjából. Kérdés, hogy az intelligens világháló lesz-e az új digitális könyvtári világháló kialakulásának motorja. (Joint 2008)

A szemantikus web és a digitális könyvtárak témakörben a legújabb áttekintés (Burke 2009) kísérleti alkalmazásokat mutat be, és az alábbi következtetéseket fogalmazza meg. A szemantikus web izgalmas bővítési potenciált jelent a webes könyvtári szolgáltatások terén, ugyanakkor óriási idő- és erőforrásigénye van a metaadatok és a taxonómiák készítésének és fenntartásának. A jövő legnagyobb kihívása a metaadatok és az ontológiák automatikus generálása. Egyre több alkalmazás használja a szemantikus világhálóra vonatkozó ajánlásokat. Teljes szemantikusweb-alkalmazás még nem készült, mert a könyvtárak ragaszkodnak meglévő rendszereikhez. A könyvtárosoknak amellet, hogy jelentős szakértelemmel rendelkeznek az intelligens világháló fejlesztésének kulcsterületein, különösen a katalogizálásban és az osztályozásban, értik is az információs erőforrások komplexitását, tapasztalattal bírnak a szabványok és az ellenőrzött szótárak fejlesztésében is. A könyvtárosok erősen elkötelezettek az információhoz való szabad hozzáférésnek és a használók személyiségi jogainak védelme területén.

A „szabványosítás” folyamata a szemantikusweb-fejlesztések során

Az intelligens világháló megvalósulásának legfontosabb előfeltétele, hogy a tartalmak létrehozói a fejlesztések során tartsák magukat az általánosan elfogadott szabványokhoz, illetve ajánlásokhoz. A World Wide Web konzorciumot (W3C) 1994-ben azért hozta létre Tim Berners-Lee, hogy elősegítse a webtechnológiák összhangját, előmozdítsa a világháló fejlődését, és garantálja széles körű felhasználhatóságát.

A W3C tevékenysége az alábbi hét alapelve épül:

- ☉ A webszabványoknak nyíltaknak kell lenniük!
- ☉ Web mindenkinek! (Tekintet nélkül az anyanyelvre, eszközkészletre)
- ☉ Web minden eszközön! (A számítógépeken kívül PDA-n, mobiltelefonon, televízión)
- ☉ A webtechnológiák legyenek együttműködők! (A webtechnológiák széles körére épít, hiszen nincs olyan technológia, amely valamennyi igényt ki tudna elégíteni.)
- ☉ Tudásalapú támogatott keresés és megosztás. (Az információ mind az emberek, mind a gépek számára legyen feldolgozható!)
- ☉ Bizalom! (Az együttműködéshez szükséges technológiák)
- ☉ A web legyen biztonságos, kiszámítható és bizalmi! (Pataki, Richter 2008)

A szemantikusweb-fejlesztéseknek széles társadalmi megegyezésen kell alapulniuk. Ennek érdekében széles szakértői és érdeklődői csoportokat kell bevonni a fejlesztésekbe. A szemantikusweb-szabványok fejlesztésében szakmailag kompetens személyek vesznek részt. A társadalmi megegyezésben tehát a tudományos szempontból már korábban megszerzett tekintély, színvonal szerepet kap.

A társadalmi ellenőrzés, az összes szereplő bevonásának gyakorlatát jól jellemzi a szabványosítás folyamata. A szabványokat a W3C konzorcium tagjaiból álló munkacsoportok készítik.

Az első dokumentum a *munkaterv* (working draft), amelyet közzétesznek, így bárki hozzászólhat. A jelentős változtatások után rendszerint többször is kiadják, figyelembe véve a javasolt módosításokat. Amikor már nagyobb változtatásokat nem szeretnének a résztvevők és hozzászólók, akkor publikálják az *előzetes javaslatot* (candidate recommendation). Ez alapján készülnek el az első implementációk, gyakorlati alkalmazások. Ezt követi a *javaslat* (proposed recommendation), amelyet már széles körben használnak. W3C ajánlássá akkor válik a dokumentum, amikor a tagok és az igazgató is elfogadják. Erről már elmondható, hogy a webes társadalom széles körben elfogadta, és *de facto* szabványnak tekinti. A hosszú procedura, a javaslatok nyilvános vitája nélkül nem lehetne elérni, hogy a konzorcium keretei között valóban széles körben elfogadott ajánlások szülessenek.

A reprezentációk

Ahhoz, hogy a valóságot le lehessen írni, először a dolgokat kell számítógéppel is dekódolhatóan megnevezni. Erre szolgál az URI (Uniform Resource Identifier = egységes erőforrás-azonosító) a szemantikus web modelljében. Az URI-k szabványa 1998 augusztusában jelent meg. A legjellemzőbb tulajdonságai a szabvány alapján a következők:

1. Egységes (Uniform):

- Különböző típusú forrásazonosítókat is megenged ugyanabban a kontextusban, abban az esetben is, amikor a szóban forgó forrásokhoz való hozzáférés különböző módon lehetséges.
- Megengedi a különböző típusú forrásazonosítókon keresztül a közös szintaxisnak az egységes szemantikai interpretációját.
- Új forrásazonosítók vezethetők be úgy, hogy nem hatnak zavaróan a létezőkre.
- Megengedi, hogy az azonosítók a legkülönbözőbb kontextusokban használhatók legyenek,

így azt is, hogy újabb alkalmazások, protokollok felülírják a jelenleg létező, széles körben használt forrásazonosítókat.

2. Erőforrás (Resource)

- Bármilyen lehet forrás, ami egyedileg azonosítható (ld. a fentebb felsorolt példákat, és vö. Szeredi, Lukácsy, Benkő 2005; Fülöp, Kovács, Micsik 2005);
- Az erőforrások vagy azok csoportjai lényegében a dolgok fogalmi felosztásából eredeztethetők, és nem kötődnek szigorúan azon egyedi tárgyhoz, ami éppen egy időben az erőforrás mögött megtalálható. Egy állandó erőforrás-azonosító mögött lévő tartalom aktuálisan változhat, ha fogalmi szinten ugyanarra vonatkozik (pl. Budapest mai időjárás-jelentése).

3. Azonosító (Identifier)

- Az azonosító egy olyan tárgy, amely azonosíthatósággal rendelkező dolgokra vonatkozik. (Berners-Lee 2005) Pepper (2008a) a tématerképeket és a szemantikusweb-technológiákat egyaránt identitás-alapúaknak nevezi, hiszen a világ létezőinek egyértelmű azonosítása és hozzá egyértelmű referencia megadása képezi az alapját az adatok hálózatba kapcsolásának.

Minden erőforrásnak, fogalomnak, dolognak, tárgynak rendelkeznie kell URI-val, ha arról állításokat kívánunk tenni. A szemantikus web későbbi szintjei URI-közötti kapcsolatokat definiálnak.

Szintaktikai szint

A világhálón lévő dolgok azonosítása után szükség van egy közös szintaxisra, ez az XML (eXtensible Markup Language = Bővíthető leíró nyelv) általános célú leíró nyelv. „*Ahogy a HTML lehetővé tette, hogy minden számítógépfelhasználó olvashassa az internetes dokumentumokat, úgy az XML – az inkompatibilis számítógépek bábeli zűrzavarában is – az eszperantót jelenti, melyen mindenki olvashat és írhat.*” (Salgáné 2004) Az XML először 1998-ban jelent

meg W3C ajánlasként. (Bray, Paoli, Sperberg-McQueen 1998), a legfrissebb, az ötödik kiadás pedig 2008 novemberében. (Bray, Paoli, Sperberg-McQueen, Maler, Yergeau 2008)

Egy XML dokumentum szöveges állomány, amelyet arra terveztek, hogy strukturált formában legyen képes szövegeket tárolni. Az XML lényege az egységes szintaxis, amely ugyanakkor nagy szabadságot enged egy-egy új dokumentum létrehozása során. (Szeredi, Lukácsy, Benkő 2005)

A HTML-hez (Hypertext Markup Language = Hiperszöveg leíró nyelv) hasonlóan elemeket és attribútumokat használ. Az XML bővíthetősége abban rejlik, hogy míg a HTML előre definiálja, hogy az egyes elemek és attribútumok mit jelentenek, itt azok csak az egyes adatsoportok elválasztására szolgálnak, illetve a nyelv az adatok értelmezését meghagyja az alkalmazásoknak.

Az XML-lel kapcsolatos tevékenység nagyon kiterjedt a W3C-n belül. Jelenleg tíz önálló munkabizottság foglalkozik XML fejlesztések különböző aspektusaival.

Az XML legszembevetőbb problémája, hogy mivel az egyes elemek elnevezése teljesen szabad, két dokumentum egybeszerkesztése lehetetlen lenne. Erre a névtér mechanizmusok kínálnak megoldást. Egy egyedi XML dokumentum, alkalmazás során használt egyedi elem- és attribútumnevek a lokális nevek, amelyek ütközhetnek két alkalmazás összekapcsolása esetén. Például elképzelhető, hogy az <m> elemnév alatt az egyik könyvtár a „munkatársakat” érti, a másik pedig a „műveket”. Ekkor az egyedi nevek kaphatnak egy prefix elemet, amelynek használatával univerzális vagy minősített nevekké tehetőek, amelyek már nagy valószínűséggel nem ütköznek és a két könyvtár adatbázisa összekapcsolható. Pl. a fenti példát használva prefix lehet a következő: <n :m> vagy: <s :m>. (Szeredi, Lukácsy, Benkő 2005) A névtérekre vonatkozóan jelenleg a 2009. decemberben megjelent ajánlás a legfrissebb. (Bray, Hollander, Layman, Tobin, Thompson 2009)

Az XML nem specifikus a szemantikus webre, önállóan is számos alkalmazási területe lehetsé-

ges. 2004 után az XML nagy teret kapott a hazai könyvtári szakirodalomban is. 1999. május 19-én volt az első nagy értekezlet az SGML és az XML nyelvekről, ahol az előadók ismertették a szintaxisokban rejlő lehetőségeket, illetve konkrét alkalmazásokat mutattak be. Az előadások elsősorban az SGML könyvtári alkalmazásaira fókuszáltak. (Moldován 1999) Az XML könyvtári rekordok formai leírásában való használhatóságát kutatták a Debreceni Egyetem Komputergrafikai és Könyvtárinformatikai Tanszékén. A projekt azzal a céllal indult, hogy lehetővé váljon a bibliográfiai leírás formalizálása az XML jelölőnyelv segítségével. (Salgáné 2004, Salgáné 2005) A HunTéka könyvtári rendszerben, annak megjelenésétől kezdve több funkcióra is XML-t használnak (konfigurációs beállítások leírása, felhasználói felületek leírása, Z39.50 adat-transzformáció, webes lekérdező motor egyes feladatainak megoldása, Tóth K. 2003) *Bíró Szabolcs* (2004, 2005) az XML digitalizálásban való használhatóságát emeli ki, tekintettel az időtállóságra. Az XML a tartalomra koncentrálnak szemben a HTML-lel, amelyet a megjelenítésre fejlesztettek ki. A nyelv a digitalizált állományok későbbi nemzedékekre való átörökítésében játszik szerepet. Az XML-nek könyvtár-informatikai szempontból a legfőbb erénye a szöveges állományok kezelésében való széles körű alkalmazhatóság.

Szemantikai szint

Az intelligens világháló legfontosabb láncszeme az RDF (Resource Description Framework = Erőforrás Leíró Keretrendszer), amellyel erőforrások írhatók le, méghozzá egymáshoz képesti viszonyaik megjelenítésével. Az RDF-logikán alapuló alkalmazások önállóan is megállják a helyüket, könyvtári környezetben egyre szélesebb körben kezdenek elterjedni. (Néhány a példák közül: Campbell, Fast 2004; Han 2005; Hecht, Haslhofer 2005; Nucci, Barbera, Morbidoni, Hahn, 2008; Malmsten 2008; Burke 2009) Az RDF egyben a szemantikus web legrégebbi

specifikációja, amelynek első generációja még a kilencvenes évekből származik, a legutóbbi ajánlás pedig 2004. február 10-én jelent meg. (Klyne, Carroll 2004) Ez a verzió elérhető magyar fordításban is a World Wide Web Konzorcium magyar irodájának oldalán. (Pataki E. 2004a)

Ez összesen hat dokumentumot jelent:

- az RDF/XML szintaxis specifikációja;
- az RDF szókészlet leíró nyelv 1.0: RDF séma (ennek ismertetését ld. később);
- az RDF bevezető tankönyve;
- az RDF alapfogalmai és absztrakt szintaxisa;
- az RDF szemantikája;
- az RDF tesztsorozata.¹ (Pataki M. 2005)

Az RDF-fel bármiről lehet állításokat tenni, ami azonosítható a weben. A szintaxis egységes keretet biztosít az információtartalom leírására, lehetőséget biztosít az információcserére, így – az URI-kkal egyértelműen azonosított fogalmak közötti relációk definiálásával – az RDF-fel létrejöhethet az alkalmazások határain átnyúló tudásreprezentáció.

Az RDF adatmodellben négy halmazt definiáltak:

1. Erőforrások – ez a halmaz bármi előforduló entitást tartalmazhat, aminek van URI-ja. Ez az összes olyan dolog halmaza, amelyre RDF kijelentés vonatkozhat.
2. Tulajdonságok – az erőforrásokhoz kapcsolódó jellemzők, amelyeknek szintén van URI-ja. A tulajdonságok lehetnek erőforrások is, tehát részalmazát képezik az előbbinek. Minden tulajdonságnak van jelentése, meghatározható, hogy milyen erőforráshoz kapcsolható, és milyen értéket vehet fel, valamint hogy milyen viszonyban van más tu-

lajdonságokkal.

3. Literálok – karaktersorozatok.
4. Kijelentések – ez utóbbiak alany, állítmány és tárgy kapcsolatok. Mindhárom elem egy-egy URI: az alany tetszőleges RDF-erőforrás, az állítmány tetszőleges RDF-tulajdonság, a tárgy pedig tetszőleges RDF-tulajdonság vagy literál lehet. (Szeredi, Lukácsy, Benkő 2005)

A 2006 májusában a 15. nemzetközi World Wide Web konferencián, Edinburgh-ban bemutatott szemantikus Wikipédia például a következő RDF-kijelentéseket tartalmazta (Völkel, Krötzsch, Vrandecic, Haller, Studer 2006):

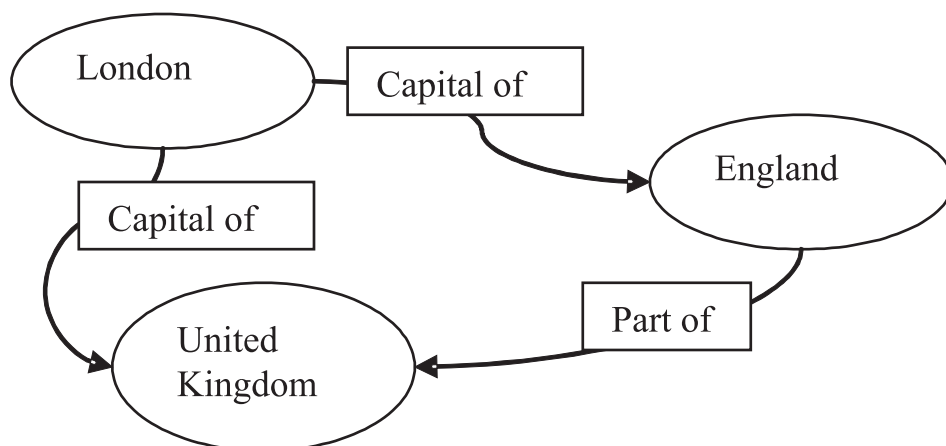
London is the capital of the United Kingdom (London az Egyesült Királyság fővárosa)

London is the capital of England (London Anglia fővárosa)

England is a part of the United Kingdom (Anglia az Egyesült Királyság része)

„London”, az „Egyesült Királyság” és „Anglia” a kijelentésekben erőforrások, és rendelkeznek egy-egy URI-val. A „része” és a „fővárosa” esetünkben tulajdonságok, amelyek szintén rendelkeznek URI-val. Anglia az egyik kijelentésben tárgyként, a másikban alanyként szerepel.

Az RDF-et szemléletesen tekinthetjük címkézett gráfoknak, amelyek kiindulási pontja egy alany, a végpontja egy tárgy, a gráf éle pedig az állítmány. A fenti kijelentések a következőképpen jeleníthetők meg címkézett gráfok segítségével:

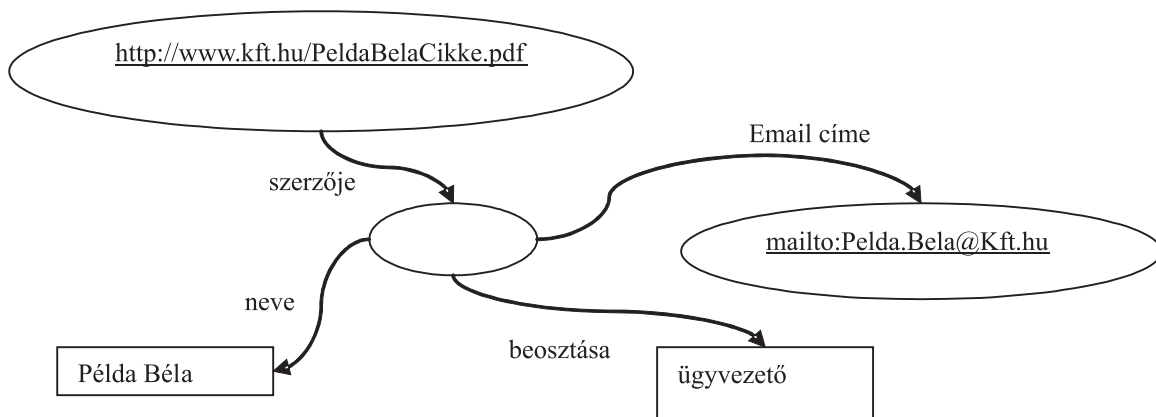


Ha két, egymástól különböző hármasban azonos URI-k találhatóak, akkor azokat a gép azonosítja, hogy ugyanazon dologról szólnak. A leírást így tovább folytathatnánk London vagy Anglia lakosságának számával, az Egyesült Királyság többi részével stb.

Az RDF-szabvány tartalmaz egy XML alapú szintaxist is. Az RDF-adat tehát megjelenhet hármasok halmazaként, mint címkézett, irányított gráf és mint XML-adat. Az RDF-elmélet szempontjából a gráfmodell az elsődleges fontosságú, a hordozhatóság és a gépi reprezentáció oldaláról viszont az XML-reprezentáció az igazán releváns. (Szeredi, Lukácsy, Benkő 2005) Az információ RDF-es reprezentációját nagyon sok alkalmazás használja önállóan a szemanti-

kus web többi rétegétől függetlenül is. Ilyenek a később részletesen ismertetésre kerülő alkalmazások közül az Arizonai Egyetem által kifejlesztett RDF alapú digitális könyvtári rendszer (Han 2005) vagy a *BRICKS* projekt. (Hecht, Haslhofer 2005)

Az RDF lehetőségeit nagyban kitérítve az üres csomópontok használatának lehetősége. Ezek használata során nem kell a triplet valamennyi elemének egy konkrétan megnevezett dolognak lenni, így például azt a kijelentést, hogy egy meghatározott URI-val azonosított cikk szerzőjének neve *Példa Béla*, beosztása *ügyvezető*, e-mail-címe pedig *Pelda.Bela@Kft.hu* a következő gráfokkal lehet reprezentálni:



Az RDF legfőbb erénye az egyszerűség. Ezzel a legkülönbözőbb tartalmak formalizálására felhasználható. A dolgozatnak nem célja a szintaxis részletes ismertetése, ezért ezen a ponton meg kell állnunk az RDF logika alapvető tulajdonságainak bemutatásával. (Részletes ismertetését ld. Passin 2004; Gottdank 2005; Szeredi, Lukácsy, Benkő 2005)

A szemantikus web tudásszervezési rendszerei

Az intelligens világháló program célja egy következtetésekre is alkalmas globális hálózati metaadat-infrastruktúra megalkotása. Szeredi,

Lukácsy, Benkő (2005) szemléletes példáját idézve megérthetjük, hogy ehhez az egyszerű RDF-állításokhoz képest milyen további fejlesztések szükségesek. „*RDF segítségével képesek vagyunk leírni azt, hogy egy ember barátja egy másiknak. Ezt úgy tehetjük meg, hogy konstruálunk egy RDF kijelentést, amelynek alanya és tárgya két ember, predikátuma egy olyan erőforrás, amelyről tudjuk, hogy a „barátja” viszonyt azonosítja valahogy. (...) Kérdés ezek után, hogy el tudjuk-e dönteni, hogy az egyik ember ismeri-e a másikat? A válasz evidens, hiszen józan paraszti eszünkkel tudjuk, hogy egy ember nyilván ismeri a saját barátját. Vegyük észre ugyanakkor, hogy valójában ez egy olyan következtetés volt a részünkről, amelyhez háttértudásra volt szük-*

ségünk. *Mi tudjuk ugyanis, hogy valakik nem lehetnek barátok, ha nem is ismerik egymást. (...) Valami olyan kijelentésre lenne szükségünk, hogy 'bármely két erőforrás között, amelyek n:barátja viszonyban állnak egymással, fennáll az s:ismerőse viszony is. 'Ezen tudás birtokában egy gép már képes lenne kikövetkeztetni, hogy a két ember ismeri egymást, bár ez így explicit módon továbbra sem lenne leírva.'* (105–106.) A sémák és az ontológiák a következtetésekhez szükséges háttértudást tartalmazzák.

A sémák és ontológiák szerepe az, hogy a háttértudás reprezentálása is egységes keretek között, a különböző alkalmazások által olvasható módon történjen. *„Elképzelhető, hogy két adatbázis különböző azonosítókat használ azonos dolgokra. A programnak, amely össze szeretné hasonlítani, vagy kombinálni a két adatbázisból származó adatokat egymással, tudnia kell, hogy a két terminus ugyanazt jelenti. Ideális esetben a program képes felkutatni az ilyen közös jelentéseket, bármilyen adatbázissal is találkozik. Az ilyen problémákra a szemantikus web harmadik nagy komponense jelenti a megoldást, az ún. ontológiák.*” (Berners-Lee, Handler, Lassila 2001) Jóllehet az ontológia több ezer éves filozófiai terminus, a 70-es évektől erőre kapó mesterségesintelligencia-kutatásokban a szó új jelentéssel gazdagodott: ekkor a szoftverekkel generált mesterséges világok megnevezésére használták. A 90-es évektől lett jelentősége a tudásmenedzsmentben az ismeretterületek formális tudásreprezentációjának. (Ungváry 2004)

Ebben az értelemben számos definícióval találkozhatunk. Vickery (1997) szerint az ontológia *„megosztott konceptualizálás formális, explicit leírása”*. A barátságatlannak tűnő definíció mögött könnyen belátható dolgok vannak. A világ jelenségei fogalmakban képeződnek le az ember számára (konceptualizálás). Az arról szóló nézeteket, hogy mely jelenség mely fogalmi körbe tartozik, egy adott nyelvet használó közösség tagjai adják, ezért hívhatjuk ezt megosztott konceptualizálásnak. Mindezt gépek számára is olvasható formában (formálisan, formális logikai szabályokra alapozva) kell kifejeznünk,

leírunk. (Gilchrist, 2003; Vickery 1997) Ennél emészthetőbb Ungváry (2004) definíciója, amely szerint az ontológia egy meghatározott ismeretterület formális modellje. Némi átfedés van a taxonómiák és a teauruszok fogalmi körével. A magyar szabvány szerint a teaurusz *„természetes nyelven kifejezett fogalmak olyan tartalmilag szabályozott, szükség szerint változtatható szótára, amelyben feltüntetik a legfontosabb fogalmi összefüggéseket. A teaurusz fő rendeltetése információk feldolgozása és keresése. A teauruszban a fogalmak lexikai egységek formájában jelennek meg.”* (MSZ 3418 1987) Ezt egészítette ki Horváth Tibor és Sütő Péter (2001) azzal, hogy a teauruszok valamely szakterülethez tartoznak.

A másik, ontológiák szempontjából jelentős átfedésekkel rendelkező terminus a taxonómia.

„A taxonómia olyan ellenőrzött szótár (egyben osztályozási rendszer), amely a dokumentumokat, illetve egyéb információforrásokat az általuk képviselt fogalmak hasonlósága alapján rendezi és csoportosítja – automatikus osztályozás esetén klaszteres, nyelvi analízis, illetve numerikus, statisztikai, vagy kombinált módszerek alapján, viszonylag alacsony hierarchiaszinten –, ezért a számítógépes kezelése áttekinthetőbbnek tűnik, mint a bonyolultabb struktúrákat képviselő osztályozási rendszereké.” (Horváth 2006:212.)

Tim Berners-Lee megfogalmazásában az ontológiák lényegében speciális taxonómiák, amelyek amellet, hogy megmutatják a fogalmak egymáshoz képesti fogalmi hierarchiáját, a formális logikai szabályok megjelenítésével még többre képesek az emberi gondolkodás reprezentálásában. (Berners-Lee, Handler, Lassila 2001) A többlet a generikus hierarchialáncokban rejlő lehetőségek automatikus kihasználásában van. Ennek lényege, hogy a tulajdonságok öröklődnek az általánostól a speciálisabb felé. (Ha a *kutya* tulajdonsága az ugatás, akkor ez a *németjuhászok* tulajdonsága is, és azon belül minden egyes egyedé, így pl. *Tappancs* is. Vö. Ungváry 2004) A magyar nyelvű ontológiafejlesztések egyik nagy ívű projektje az elmúlt években a

Magyar Egységes Ontológia (MEO) megvalósítását tűzte ki célul. A kutatás zárójelentése szerint „*nemhogy az ontológiák tényleges ipari alkalmazásától messze vagyunk, de ma még mindenhol az alapok kialakításánál, megszilárdításánál tartanak.*” (Szakadát 2006) A tudásrepresentáció korlátozott lehetőségei miatt tehát egy következtetésekre is alkalmas globális metaadat-infrastruktúra megvalósításától még igen messze járunk.

A szemantikusweb-fejlesztések számos lehetőséget kínálnak az háttértudás formális leírására. Kifejezések, dolgok osztályait lehet meghatározni az RDF-séma segítségével. A legfrissebb RDF-séma W3C ajánlás 2004. február 10-én jelent meg. (Brickley, Guha 2004) Az adatmodellje segítségével már képesek lehetünk egyszerűbb következtetések levonására, a bonyolultabb logikai műveletek azonban megkövetelik az OWL (Web ontológia nyelv) használatát. Ennek előzménye a DAML+OIL (DARPA² Agent Markup Language + Ontology Inference Layer / Ontology Interchange Language) leíró nyelv volt, amelyet az OWL teljes egészében felváltott. (Conolly et al. 2001) Az OWL esetében tudatosan cserélték fel a betűszó első két tagját, hogy a közismerten bölcs *bagoly*ra utaljon.

A specifikáció W3C ajánlasként 2004 februárjában jelent meg. (McGuinness, Van Harmelen 2004)

Ez alapján a következő magyar fordítások készültek:

- Áttekintés,
- Útmutató – amely az OWL-t egy részletes példával szemlélteti;
- Referencia – az OWL modellezési primitívek jól szerkesztett, informális leírása,
- Szemantika és absztrakt szintaxis – az OWL nyelv normatív definíciója,
- Az OWL tesztoszorozata – az OWL használatát illusztráló példák, a konst-

rukciók formális jelentése és a témák kifejtése,

- OWL alkalmazási esetek és követelmények
 - egy webontológia nyelv felhasználási forogatókönyve, céljai és követelményei. (Pataki E. 2004b)

A legfőbb különbség az RDF sémával szemben, hogy míg az csak osztályok és alosztályok definiálását teszi lehetővé, az OWL az összes halmazelméleti műveletre képes. Osztályok konstruálhatók megnevezéssel, az egyedek felsorolásával, tulajdonságkorlátozással, osztályok uniójaként, metszeteként, komplementeként; lehetséges az osztályok egyedeinek felsorolása, az osztályokba tartozó egyedek tulajdonságaik szerinti definiálása. Az OWL a tulajdonságok terén is magasabb szintjét képes megragadni a gondolkodásnak; általa tulajdonságokat lehet szimmetrikusnak, tranzitívnek, függvénynek stb. deklarálni. Az OWL esetében megragadható az osztályok, tulajdonságok vagy egyedek ekvivalenciájának, diszjunktságának problémája. Kezelhető a lokális és globális szárosság-megkövetések vagy a kardinalitás problémája (Pataki E 2004b; Gottdank 2005; Szeredi, Lukácsy, Benkő 2005; Herman 2006).

Az OWL valamennyi tulajdonságot egy-egy arra a célra rendelt attribútummal fejezi ki. A dolgozatnak nem célja ezen attribútumok bemutatása, így csak egyetlen példával illusztráljuk az OWL szintaxist.

```
<owl:Class>
  <owl:intersectionOf rdf:parseType="Collection">
    <owl:Class rdf:about="#Szociológus"/>
    <owl:class>
      <owl:oneOf rdf:parseType="Collection">
        <owl:Thing rdf:about="Amberg Eszter" />
        <owl:Thing rdf:about="Kis Klára" />
        <owl:Thing rdf:about="Marosán Ágnes" />
        <owl:Thing rdf:about="Péterfi Rita" />
        <owl:Thing rdf:about="Fehér Miklós" />
        <owl:Thing rdf:about="Ládi László" />
        <owl:Thing rdf:about="Somogyi József" />
        <owl:Thing rdf:about="Tóth Máté" />
        <owl:Thing rdf:about="Vidra Szabó Ferenc" />
      </owl:oneOf>
    </owl:class>
  </owl:intersectionOf>
</owl:Class>
```

A leírás egy metszetkonstrukció, amely a *Könyvtári Intézet Kutatási és szervezettefejlesztési osztályának* azon munkatársait írja le, akik szociológusok. Ehhez meg kellett alkotnunk megnevezéssel a *szociológusok* osztályát, majd az egyedek felsorolásával a Kutatási és szervezettefejlesztési osztály munkatársainak osztályát. Végül a metszetképzés halmazelméleti műveletével definiáltuk azokat az egyedeket, amelyek mindkét halmaz elemeinek tekinthetők (lásd az előző oldali ábrát)

Az OWL-nak három növekvő erejű kifejező alnyelve van a különböző felhasználási igényekhez igazítva:

- OWL Lite – osztályozási hierarchiakat és egyszerű korlátozásokat alkalmazó felhasználók támogatására készült.
- OWL DL – a maximális kifejezőképességet igénylő felhasználóknak készült. Az OWL DL-ben a számíthatóság és az eldönthetőség is megmarad. Az összes nyelvi konstrukciót tartalmazza, de azok használhatósága csak bizonyos korlátozásokkal lehetséges. A DL (Description Logic) a leíró logikákra utal.
- OWL Full – olyan maximális kifejezőképességet és az RDF teljes szintaktikai szabadságát igénylő használóknak készült, akik lemondanak a kiszámíthatósági garanciákról. (Kérdés, hogy a szoftvertámogatás – éppen a túlságosan gazdag logikai struktúrák következtében – megvalósítható-e.) (Pataki E. 2004b)

Az alnyelvek mindegyike az egyszerűbbnek a kiterjesztése. Így például minden érvényes OWL Lite következtetés egyben érvényes OWL DL következtetés is, és minden érvényes OWL DL következtetés egyben érvényes OWL Full következtetés is. Minden legális OWL Lite ontológia egyben legális OWL DL ontológia is, és minden legális OWL DL ontológia egyben legális OWL Full ontológia is. (Pataki E. 2004b)

A használóknak maguknak kell eldönteniük, melyik alnyelv a legalkalmasabb a céljaikra. (Digitális könyvtári környezetre való alkalmazhatóság tekintetében beszédes tény, hogy a DL mögött egyesek a Digital Library kezdetűit vélik felfedezni. Vö. Krause 2008)

Az OWL az általa kifejezett ontológiák segítségével az emberi tudásnak egy magasabb szintjéig képes reprezentálni és alkalmazni gépi műveletek végrehajtásához. Háttértudásként az ontológiáknál már nem csupán az osztály–alosztály relációk vagy a tulajdonságok hierarchikus rendje van jelen, hanem a formális logikai szabályokat is alapul veszik a következtetésekhez. Az OWL gyengéje ugyanaz, mint ami az ereje: kérdéses ugyanis, hogy a szigorú logikai struktúrákkal hatékonyan modellezhető-e a világ létezői és a közöttük lévő relációk. A világ létezőinek formális leírása szükségszerűen egyszerűsítésekkel jár. A radikális egyszerűsítésekkel olyannyira elnagyoltta válik a világról alkotott kép, hogy nem leszünk képesek érvényes következtetéseket levonni. Ha pedig túlságosan részletesen próbáljuk meg leírni a valóságot, akkor maga az ontológia válik használhatatlanná a bonyolult relációk kuszasága miatt.

Könnyen belátható, hogy óriási munka lenne a világ minden létezőjét ontológiákban rögzíteni. Az is világos, hogy nagyon sok ismeretet leírtak már formálisan más tudásszervezési rendszerekben (tezauruszokban, taxonómiákban, folkszonómiákban, tárgyszójegyzékekben). Szükség volt egy olyan szabványos módszerre, amellyel lehetővé válik ez utóbbiak exportálása a szemantikus webbel kompatibilis formátumba. Erre szolgál a SKOS (Simple Knowledge Organization Systems). Fejlesztése viszonylag későn kezdődött az intelligens világháló többi technológiájához képest, ugyanakkor maga a szintaxis olyannyira ígéretes volt, hogy bő két év alatt a SKOS szemantikusweb-ajánlássá nőtte ki magát. Az első munkaterv 2007. május 16-án jelent meg, (Isaac, Phipps, Rubin 2007) egy 2009. június 15-i javaslattevő után (Miles, Bechhofer 2009a) ugyanazon év augusztusában megjelent a W3C ajánlás (Miles, Bechhofer 2009b).

A SKOS ereje az egyszerűségében rejlik, hiszen nem mindig a legnagyobb szemantikai kifejezőerővel bíró megoldás a leghatékonyabb. Egy szemléletes példát idézve: „*amíg az OWL egy nehéz kalapács, addig a SKOS egy egyszerű diótörő gép, illetve amíg az OWL egy Harley*

Davidson, addig a SKOS egy sima kerékpár.” A SKOS az OWL-lal együttműködhet, így optimalizálható a tudásreprezentáció mélysége. (Krause 2008)

A digitális könyvtári alkalmazás szempontjából a SKOS közvetlen eredményeket jelenthet a meglévő tudásszervezési rendszerek és a szemantikus web integrálása terén. A könyvtárostársadalom a fejlesztések kezdeti stádiumától kezdve figyelemmel kísérte a SKOS-t. (Cantara 2006) Lényegében a SKOS-nak az a célja, hogy a könyvtári osztályozást és a szemantikusweb-technológiákat egységes keretek közé integrálja. (Krause 2008) A könyvtárak már az intelligens világhálót megelőzően is nagy mennyiségű szemantikailag kódolt adattal rendelkeztek. A könyvtári osztályozási rendszerek évezredek tapasztalatait ez a specifikáció fogja a szélesebb webes közösség számára elérhetővé és használhatóvá tenni.

A lekérdezés szabványai

A tudás formális leírását követően szükség van az ismeretek kinyerésére is. Mivel a szemantikus web valamennyi specifikációja az RDF logikát követi, a lekeresésnek is tripletek mentén kell történnie. A szemantikus web lekérdező nyelve a SPARQL (rekurzív akronímia: SPARQL Protocol And RDF Query Language).

Az RDF Adathozzáférési Munkacsoportja (Data Access Working Group = DAWG) a W3C égisze alatt jött létre 2004 februárjában azzal a céllal, hogy az RDF címkézett irányított gráfjaihoz készítsen egy nyelvet, amely különböző adatforrásokon keresztül képes lekeresni a tárolt információt, függetlenül attól, hogy a tartalom RDF-ben készült, vagy csak úgy lett megjelenítve. Az első munkaterv 2004 októberében jelent meg, a legfrissebb SPARQL W3C ajánlás pedig 2008. január 15-én. (Prud'hommeaux, Seaborne 2008) A nyelv logikája nagyon egyszerű, egy gráfmintát tartalmaz, amelyben az alany, állítmány vagy a tárgy helyén állhatnak változók. Ide próbálja meg behelyettesíteni a hármas halmazban lévő

URI-kat, és ellenőrzi, hogy az így kapott gráf az eredeti RDF gráf részgráfja-e. (Dodds 2005)

A szemantikus web könyvtári alkalmazásai

A könyvtári alkalmazások közül az első kísérletek az XML és az RDF implementálásával kezdődtek. *Yan Han* az Arizonai Egyetem Könyvtárában egy RDF-alapú kísérleti digitális könyvtári alkalmazást fejlesztett ki. A munka 2005 áprilisában indult. A kísérleti alkalmazásban az RDF technológiát használták a tudásmenedzsment céljaira és metaadatok kezelésére. A kutatás motivációját az ismeretterületek közötti interoperabilitáson alapuló információcserére vonatkozó igény jelentette. A kísérleti alkalmazás egy RDF-en alapuló folyóirat-kezelő rendszer volt. Ez később kiegészült egy elektronikus disszertáció- és szakdolgozat-kezelő alkalmazással, amely ugyanazon a rendszeren belül futott. (Han 2005) A kutatás igazolta az RDF digitális könyvtári területen való használhatóságát, de nem eredményezett a napi gyakorlatban is működni képes kész rendszert.

Mary Burke (2009) vázlatos, de majdnem teljes áttekintést adott a digitális könyvtári szemantikusweb-fejlesztésekről és kutatásokról. Összegyűjtötte a legjelentősebb projekteket, amelyek a szemantikusweb-technológiák könyvtári környezetbe való adaptálását célozták. (Az alábbi példák összeállításában nagy segítségünkre volt szemléje.)

Baker (2006) a digitális könyvtárak jövőjéről szóló áttekintése szerint az elkövetkezendő időkben meghatározó trend lesz a digitális könyvtárak területén a *perszonalizáció*, a szolgáltatások testre szabhatóságának a megteremtése. Ez a trend mind a szemantikus web, mind pedig az alább tárgyalt web2.0 fejlesztésekkel összhangban van. *Nuria Ferran* és kollégái a barcelonai egyetemen erre mutattak be egy lehetséges digitális könyvtári modellt. A modell lényege, hogy a digitális könyvtári tranzakciók

valamennyi elemét (fogalmát) és a köztük fennálló relációkat leírják egy ontológiában. Ennek érdekében igyekeztek a digitális könyvtárak használatának minden elemét azonosítani és a kísérleti rendszerbe integrálni. A testre szabhatóságot két fontos tényező határozza meg. Először egy felhasználói interfészt kell kialakítani, egy olyan személyes felületet, amely tartalmazza a felhasználói profilt, a felhasználó preferenciáit és a kereséstörténetet. Másrészt a rendszernek használnia kell a használói magatartás elemzéséből gyűjtött információkat. A korai kísérleti alkalmazás végén nem jött létre kész, működő rendszer. Maga az ontológia is további fejlesztésekre szorult, viszont arra a következtetésre jutottak, hogy az ontológiafejlesztéssel létrehozott új szolgáltatások a szemantikus web szabványok alkalmazásával lesznek a meglévőkhöz illeszthetők. (Ferran, Mor, Minguillon 2005)

Campbell és Fast (2004) az egyetemi és szakkönyvtári katalógusok szempontjából vizsgálta a szemantikusweb-szabványok által kínált lehetőségeket. Véleményük szerint a katalógusrekord információtartalmát úgy lehetne bővíteni, ha egy szoftver RDF-dokumentumokat gyűjtene a világhálóról, és azokat hozzárendelné az egyes tételekhez; a művekhez a műfajt, történelmi adalékokat, további tartalmakat a szerzőről, valamint bibliográfiai jellemzőket lehetne kapcsolni. Ennek érdekében a katalogizálóknak RDF-dokumentumokat kellene keresniük, azokat megbízhatóság és hasznosság szempontjából értékelni, illetve elkezdni a könyvtári katalógusok tartalmának RDF-ben való leírását és elérhetővé tételét. A szerzők szerint az egyetemi könyvtárak katalógusai az irodalom feltárásától a weben található tartalmak értékelésének irányába fognak elmozdulni, ami merőben új szemléletet hoz majd az elektronikus szolgáltatások területén. A cikk öt esztendeje, a szemantikus-webfejlesztések első periódusában, az eufória jegyében született. Kétségtelen, hogy amennyiben az RDF-nek köszönhetően a katalógusok és a világháló valamennyi tartalma között megvalósulna az együttműködés, úgy a rekordok egy globális metaadat-infrastruktúrának a

részét képeznék. Ez a szemantikus kapcsolatok, közvetve a rekordokból elérhető információtartalom bővülésével járna. A szemantikusweb-kutatások megtorpanása nyomán az RDF öt év múlva sem vált a világhálón található tartalmak fő metanyelvévé, ami elsőszámú feltétele lenne az ilyen irányú fejlesztéseknek.

A szemantikusweb-szabványok digitális könyvtári környezetben való használatát kutatták a Gdanski Egyetem Központi Könyvtára és a Digitális Vállalat Kutatóintézet (Digital Enterprise Research Institute) közös *JeromeDL* projektjében. (Kruk, Czygan, Gzella 2008) A magát közösségi szemantikus digitális portálnak nevező kísérleti rendszer a szemantikus web és a Web2.0 számos ismervét tartalmazza. (<http://www.jeromedl.org/>) Digitális könyvtárként lehetőséget teremt a dokumentumok világhálón való publikálására. Különböző formátumokat támogat, továbbá tárolja, böngészhetővé és kereshetővé teszi azok bibliográfiai adatait. A rendszer lehetőséget teremt használói profil készítésére és testre szabására, tartalmi összefoglalók, jegyzetek írására, saját virtuális „könyvespolc” létrehozására, továbbá ezek másokkal való megosztására. Háromfajta (strukturális, bibliográfiai és közösségi) metaadattal írja le a feltöltött tartalmakat. A projektben a létező digitális könyvtári rendszerek tulajdonságait ötvözték a szemantikus tudásreprezentáció eszközeivel. Az rendszer a tartalmakról szóló strukturális információkat (pl. könyvfejezetek) ontológiákban tárolja. Az ontológiák a JeromeDL legalsó rétegei, amelyek a tartalmak fizikai reprezentációjához szükségesek, képesek a dokumentumok, digitális objektumok részeit, azok lelőhelyét és a részek közötti kapcsolatokat megjeleníteni. (Woroniecki, Kruk 2007; Kruk 2008a, Kruk 2008b)

A rendszer a közösségi hálózatok és a szemantikusweb-technológiáit ötvözi annak érdekében, hogy hatékonyabbá lehessen tenni a keresést a digitális könyvtárakban. *Sebastian Ryszard Kruk*, a JeromeDL projekt vezetője, *Ewelina Kruk* és *Katarzyna Stankiewicz* (2008) összehasonlítottak két digitális könyvtári rendszert: egy olyat, amely tartalmazza a szemantikus web és a

közösségi hálózati technológiákat, és egy olyat, amely nem. A kvantitatív adatgyűjtés eredményei azt mutatták, hogy az előbbivel a használók jobb minőségű tartalmakat nyertek ki a keresés során, és ezzel párhuzamosan elégedettebbek is voltak a rendszerrel. A kutatás során azt is kimutatták, hogy a szemantika megragadása során a digitális könyvtáraknak sokkal erőteljesebben kell támaszkodniuk az ún. „puha” szemantikára, azaz a használók által generált metaadatokra.

Digitális Archívumok

Egy európai digitális archívum felépítését célozta meg a *BRICKS* projekt (Building Resources for Integrated Cultural Knowledge Services), amely a metaadatok tekintetében támaszkodik a szemantikusweb-ajánlásokra. A *BRICKS* projekt részben egy nyílt és bővíthető közösséget, részben pedig egy technológiai platformot jelent. (<http://www.brickcommunity.org/>) A program háttérét az EU 6. keretprogramja (FP6) biztosítja. A projektnek az a célja, hogy az európai kulturális örökség gyűjteményeit egy nyílt forráskódú szoftver segítségével integrálja. A csatlakozó intézmények az ingyenes szoftverrel közös együttműködő platformon teszik közzé tartalmaikat, amelyek így a teljes rendszer számára elérhetővé válnak. Az intézmények ezzel részévé válnak a *BRICKS* közösségnek. (Jelenlegi tag az Osztrák Nemzeti Könyvtár és az olasz Consorzio Forma képgyűjteménye.) A szoftver egy metaadat-, egy tartalom- és egy gyűjteménykezelő alkalmazást tartalmaz. A metaadatkezelő alkalmazás RDF-alapú. A szemantikusweb-szabványokkal tehát lényegében egy egymással együttműködő metaadat-infrastruktúrát épít a *BRICKS* projekt, amely tetszőlegesen bővíthető. (Hecht, Haslhofer 2005)

A *Talia* megosztott szemantikus digitális könyvtári platform, amelyet olasz kutatók fejlesztettek ki, kifejezetten a társadalom- és bölcsészettudományi kutatások igényeihez (<http://trac.talia.discovery-project.eu/>). A digitális könyvtári jellemzők mellett tartalmaz elektronikus publi-

kálási rendszert is, amely a legkülönbözőbb formátumú szöveges, képi és video-állományokat képes kezelni. Minden *Talia*-ban található tartalmat (dokumentumot, erőforrást) stabilan meghatározott és változatlan URI azonosít, amely biztosítja a dokumentumok helyének állandóságát az információs térben. (Ennek a hivatkozások esetében van jelentősége.) A *Talia* teljes egészében az intelligens világháló technológiáin alapul, ami lehetővé teszi az együttműködést más szemantikusweb-alapú eszközökkel és alkalmazásokkal. (Nucci, Hahn, Barbera 2008) A *Talia* csak nagyon átfogó fogalmakat és általános relációkat magában foglaló ontológiát tartalmaz, viszont támogatja szakterületi ontológiák létrehozását. A szakterületi ontológiák importálhatók is a könyvtári rendszerbe. A *Talia*-t nyílt forráskódú licenc formájában terjesztik, így alacsony költséggel installálható egy digitális könyvtári és egy elektronikus publikálási rendszer. (Nucci, Barbera, Morbidoni, Hahn 2008) A rendszer önmagában is működőképes, azonban az egyes könyvtár vagy szervezet virtuális határai átjárhatóvá válnak, továbbá egy globális metaadat-infrastruktúra irányába.

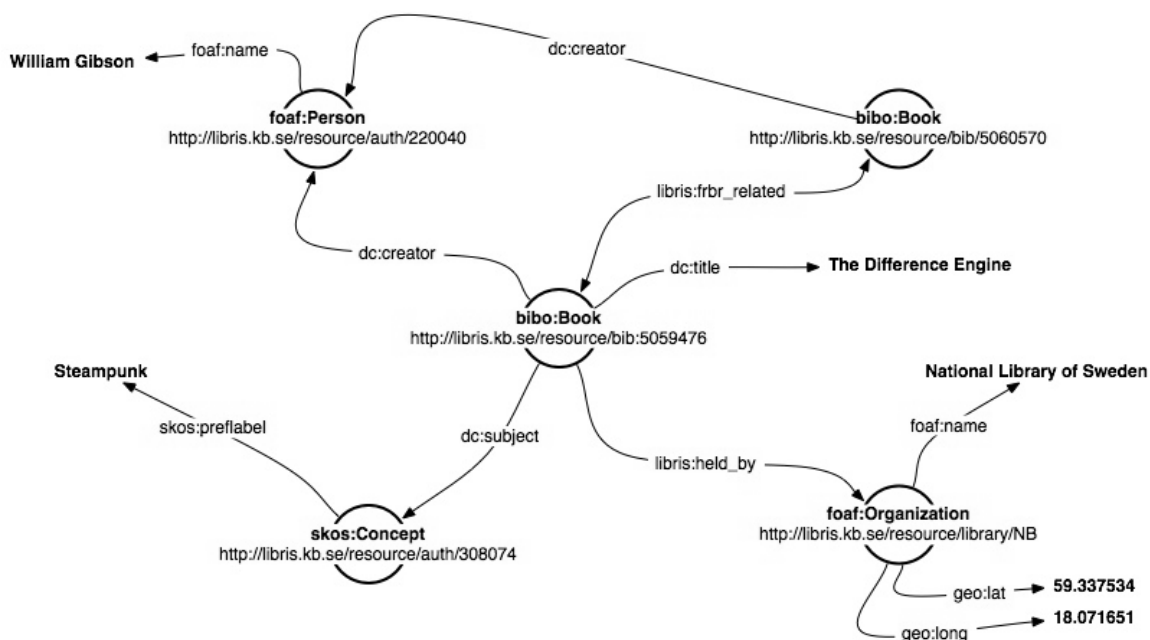
Az Amerikai Egyesült Államokban a Cornell Egyetem kutatói 1997-ben indították útjára a *FEDORA* (Flexible Extensible Digital Object Repository Architecture) projektet (<http://www.fedora.info>). A *FEDORA* digitális objektumok formális leírását tűzte ki célul: egy olyan architektúra létrehozását, amely egyaránt alkalmas digitális tartalmak tárolására, menedzselésére és hozzáférhetővé tételére. A szemantikusweb-ajánlások megjelenését követően implementálták a projektbe az RDF-et. A *FEDORA* jelenlegi 3.2 verziója RDF-tripleteket tartalmaz, amely átjárhatóságot biztosít más szemantikusweb-alapú rendszerekkel. A *FEDORA* körül – a *BRICKS* és a *Talia* platformokhoz hasonlóan – közösség van, amely amellet, hogy használja a rendszert, újabb alkalmazásokat és további eszközöket fejleszt ki hozzá. (A közösségre utal a *Fedora Commons* név.) Tekintve, hogy a *FEDORA* egy általános digitális tartalmak kezelésére alkalmas

szoftver, a használók között a könyvtárak mellett innovatív szakmai szervezetek, kutatóintézetek, múzeumok, kulturális intézmények, kormányzati szervek, vállalatok is találhatóak. A Fedorán alapul többek között a PloS (Public Library of Science), a University of Virginia Library, a National Science Digital Library és az Irish Virtual Research Library and Archive Project digitális tartalomkezelő rendszere.

Burke (2009) nem tesz említést az elmúlt évek egyik legsikeresebb könyvtári szemantikus web alkalmazásáról. A svéd nemzeti könyvtár (Kungliga Biblioteket) az eddigi példáktól eltérően nem egy teljesen új rendszert fejlesztett, hanem arra tett kísérletet, hogy az ország közös

katalógusát, a *LIBRIS*-t, amely nagy mennyiségű strukturált és ellenőrzött adatot tartalmaz, a szemantikus webhez kompatibilis módon konvertálják. A hangsúlyt az erőforrásokra mutató és az erőforrások közötti kapcsolatokra helyezték, nem pedig az egyes tételek minél részletesebb megjelenítésére. A különböző alkalmazásokkal való együttműködés és a MARC rekordok szemantikai tartalmának biztosítása érdekében a Dublin Core, a SKOS és a FOAF-elemeket kiegészítették egy saját maguk által generált ontológiában. Az adatokhoz való hozzáférést a SPARQL biztosítja. (Malmsten 2008)

Az alábbi ábra egy rekordhoz kapcsolódó elemeket mutatja be irányított gráfok segítségével. (Forrás: <http://blog.libris.kb.se/semweb/>)



Az Országos Széchényi Könyvtár a svéd nemzeti könyvtár gyakorlatához igen hasonló fejlesztéssel a katalógusában tárolt metaadatokat tette közzé, illeszkedve a létező a szemantikusweb-szabványokhoz és ellenőrzött szótárakhoz. (http://nektar.oszk.hu/wiki/Szemantikus_web)

A szemantikus web kritikája

A szemantikus web erősségét a metaadatok és a tudást reprezentáló ontológiák jelentik. A koncepció értelmében ez utóbbiak biztosítják a következtetések lehetőségét, amely nyomán valóban „intelligens webről” lehetne beszélni. Már az OWL fejlesztésekor világossá vált, hogy egy OWL Full alkalmazásban képtelenség len-

ne a világról való ismereteinket olyan módon formalizálni, hogy abban ne legyenek olyan ellentmondások, amelyek lehetetlenné teszik a szoftverekkel végzett következtetéseket. A világ dolgainak leírhatatlanságára elég a természetes nyelveket példaként hozni. A természetes nyelvek az ontológiákhoz hasonlóan megosztott konceptualizálással formalizálják a világ létezőit. (Vickery 1997) A világ létezői azonban bonyolultságuknál fogva olyannyira sokféleképpen képezhetők le, hogy elkerülhetetlenek az ellentmondások. (A nyelvi relativizmus elmélete szerint a természetes nyelvek determinálják a gondolkodást is, így a más anyanyelvűek másképpen szemlélik a világot.) A magyar csúcsontológia projekt bebizonyította, hogy egyetlen nyelv legáltalánosabb fogalmait is problémás hatékonyan megragadni. Elsősorban ennek eredményeként az ontológiák fejlesztése a szemantikus webről szóló diskurzus kezdete óta eltelt tíz évet követően még mindig gyerekcipőben jár, és nem is várható áttörés ebben a tekintetben. (Krause 2008) Nem jelent igazi megoldást a létező tudásreprezentációs eszközökhöz való fordulás sem (pl. SKOS), mert ezekkel éppen a szemantikus web legfőbb ígéretéről a formális logikákon alapuló, szoftverekkel elvégezhető következtetésekről kellene lemondani, tehát a Berners-Lee, Handler és Lassila (2001) által felvázolt lehetőségek csak nagyon szerény mértékben válnának valóra. Egyre inkább világossá válik, hogy globális metaadat-infrastruktúra építése elképzelhetetlen a tömegek munkája nélkül. A világháló szintaktikai paradigmájának sikere nagyban köszönhető az egyszerűségnek. A szabványokat felhasználói szintű számítógép-kezelői ismeretekkel rendelkező személyek is képesek alkalmazni a HTML-szerkesztő programok segítségével. A szemantikus web bonyolult szerkezete megnehezíti az alkalmazást, így ezek a szabványok nem is tudtak elterjedni. A szemantikusweb-technológiák olyan szakismeretet igényelnek, amely lehetetlenné teszi a tömeges alkalmazást. Ennek némileg ellentmond, hogy az RDF-alapú tudásreprezentáció egyre elterjedtebbé válik könyvtári körben, és mind több intézmény teszi

közé metaadatait a szemantikus webbel kompatibilis formában, összességében azonban a szemantikusweb-technológiák igen szerény mértékben vannak jelen a világhálón.

Minden létezőt magában foglaló ontológia alkalmazás hiányában a szemantikus web nem válhat globális metadat-infrastruktúrává, hanem megmarad a webes tartalmak hatékonyabb együttműködését elősegítő technológiák gyűjtőfogalmának. Az ismertetett szemantikusweb-technológiákon alapuló könyvtári alkalmazások sorát áttekintve belátható, hogy egyetlen olyan rendszer sem született, amely a lépcsős modell valamennyi szintjét magában foglalná. A legáltalánosabb az RDF, azaz a címkézett gráfok alkalmazása a metaadatok előállításában. Szakterületi ontológiák építésében történt ugyan némi előrelépés, ezek azonban nem feltétlenül biztosítják a megfelelő szintű átjárhatóságot, így a következtetések érvénye is szükségszerűen korlátozott.

Mindezidáig nem született megoldás az URI-k hatékony újrahazsnosításának elősegítésére. Egy globális metaadat-infrastruktúra alapvető feltétele, hogy azokat a dolgokat, amelyekről állításokat teszünk, alkalmazások során keresztül ugyanazzal az URI-val azonosítsuk. Ez szintén csak úgy valósítható meg, hogy ha magukat az ontológiákat is újrahazsnosítják. Amikor valaki új tartalmakat lát el szemantikusweb-technológiákhoz kompatibilis módon metaadatokkal, akkor csak ezen ontológiákból válogathatná ki, hogy milyen URI-t rendel a világ azon létezőihez, amelyekről ő (is) állításokat tesz. Eltekintve néhány olyan sikeres kezdeményezéstől, mint a FOAF, erre alig találunk példát.

A FOAF és minden más, személyek kapcsán egyértelmű azonosítóval közzétett metaadat kapcsán felmerül a személyiségi jogok kérdése. Egy nagy együttműködő metaadat-infrastruktúra alapján, amellyel különböző forrásból származó adatok összekapcsolhatók, olyan dolgok is kikövetkeztethetők lesznek gépi úton személyekről, amelyek közzétételéhez az illető nem biztos, hogy hozzájárulna. Az adatok védelme így szinte lehetetlen feladat lesz, hiszen az ösz-

szekapcsolt tartalmak, ha egyenként nem is sértenének személyiségi jogokat, együtt megjósolhatatlan, hogy milyen következtetésekre lesznek elegendők. Ha valószínűsítjük, hogy a mélyweben tárolt tartalom (a szemantikailag kódolt információ tömeggel) a szemantikus web valósággá válásával felszabadul, és hatékonyan kereshetővé válik, a probléma fokozottan fog jelentkezni.

A könyvtári szempontú áttekintés végkövetkeztetéseként csatlakozunk Szakadát István 2007-ben megfogalmazott véleményéhez. *„A szemantikus web (...) ma még sokkal inkább csak ígéret, mint valóság, és egyelőre nem is nagyon látszik, mikorra és hogyan leszünk képesek a szemantikus web program céljait megvalósítani.”* (Szakadát 2007b)

Könyvtárak a szemantikus web világában

„A könyvtárak szerepe nem változott az elmúlt évezredekben, s nem fog változni a jövőben sem: beszerezni, őrizni, hozzáférhetővé tenni a tudást és az információt, bármilyen hordozón jelenik meg, és segíteni az embereket, akik hozzá akarnak férni ezekhez.” – írta Crawford és Gorman (1995). Ha a szerepek nem is változtak, a környezet olyan új kihívásokat és lehetőségeket tárt a könyvtárak elé, amelyekkel szembenézve szükségesnek tűnik a jövőkép újradefiniálása. A környezeti változások lényegét Tim Berners-Lee egy interjúban a következőképpen foglalta össze: *„A webről szóló álmom két részből áll. Az elsőben azt látom, hogy a világháló egyre inkább képes támogatni az emberek közötti együttműködést. A másodikban az együttműködés kiterjed a számítógépekre, amelyek képessé válnak a webes adatok – tartalmak, csatolások és ember-gép tranzakciók – elemzésére.”* (Anbarasan 2000) A szemantikusweb-fejlesztések az elmúlt évek legjelentősebb lépését jelentik egy hálózati környezetben megvalósuló globális tudáscsere felé. Ha a könyvtárak továbbra is a tudás és az infor-

máció beszerzői, őrizői és szolgáltatói akarnak maradni, akkor folyamatosan figyelemmel kell kísérniük a szemantikus technológiák fejletéseinek alakulását. A web fejlődési irányainak áttekintését követően felvázolom a szemantikai paradigma megjelenése nyomán egy globális hálózati metaadat-infrastruktúra kialakításának könyvtári vonatkozásait.

A szemantikus technológiák alapját a metaadatok jelentik, a könyvtárak pedig a világ egyik legnagyobb (ha nem a legnagyobb) metaadat-kincse fölött rendelkeznek. Sőt. A könyvtárak szemantikailag kódolt metaadatokkal rendelkeznek. A „hagyományos” világháló paradigmájában a digitális könyvtárak jelentették a web azon szeletét, amelyben az információ a szemantikai tartalomnak megfelelően került elhelyezésre. Ez óriási potenciált és versenyelőnyt jelent számunkra. Mindez ki is jelöli a könyvtárak első és legfontosabb feladatát: a birtokukban lévő metaadat-tartalmakat közzé kell tenniük a szemantikus technológiákkal kompatibilis szabványok szerint. (Coyle 2008; Malmsten 2008; Burke 2009) Arra vonatkozóan kiterjedt kutatások vannak, hogy miként biztosítható a könyvtári és a szemantikusweb-szabványok közötti átjárhatóság.

A könyvtár régóta hangoztatja, hogy nem csupán dokumentumokat, hanem információt, illetve tudást szolgáltat. A tudásreprezentáció azonban csak közvetetten, a dokumentumok tartalmi feldolgozásával valósult meg. A szemantikus technológiák megadták az esélyt arra, hogy a dokumentumok szintjénél beljebb hatoljunk, és ténylegesen megragadjuk a tudást. Eddig a szemantikailag kódolt információ annyira volt elegendő, hogy releváns dokumentumokat találjunk, a szemantikus technológiákkal ezek magához a válaszhoz vezetnek el. Ehhez azonban nem elegendők a könyvtárak által közzétett metaadatok, az internetezők millióinak a munkájára is szükség van. A könyvtár nem vállalhatja fel, hogy a világ tudását rendszerezze szemantikailag kódolt formában, így egyértelmű, hogy a koordináló, „tudásmérnöki” funkciót kell betölteni a tartalmak publikálása tekintetében.

A tudásmérnöki szerepnek az új paradigmában a következőkre kell kiterjedni:

- a fogalmak egyértelmű azonosítása;
- a szabványok használatának elősegítése;
- tudásszervezési rendszerek közzététele és karbantartása.

A fogalmak egyértelmű azonosítása tekintetében alig van előrelépés. A szemantikusweb-fejlesztések területén a dolgok URI-kkal való azonosítása nem jelent egyértelmű identitást a fogalmaknak, amelyekről szó van. A könyvtárak ezzel szemben évszázadok óta rendelkeznek ellenőrzött szótárakkal, komplett szakterminológiák tudásszervezési rendszerekkel való reprezentációival. Óriási előrelépést jelentene, ha a nagy könyvtárak tárgyszórendszerei (pl. az Egyesült Államokban használt Library of Congress Subject Headings) a bennük lévő szemantikai tartalommal együtt megjelenének a FOAF-hoz hasonló egyszerű szótárakban. Ennek a lehetőséget teremtette meg a SKOS információtechnológiai oldalról. Hogy a lehetőséggel élni tudjunk, a másik oldalon, a könyvtárosoknak is lépniük kell. Sokkal intenzívebb párbeszédet kellene folytatni a szemantikus web fejlesztőivel, és bizonyítani, hogy a tudásszervezési rendszerek készítése és közzététele szempontjából a könyvtár megkerülhetetlen intézmény. A könyvtártudomány a szemantikusweb-fejlesztések kulcsterülete. (Ungváry 2004; Horváth 2006; Tóth 2007; Burke 2009) Az előrelépéshez a legjelentősebb lépést az jelenti, ha a könyvtárakban explicit formában rendelkezésre álló szellemi vagyont a szemantikus technológiák számára is felszabadítják, közkinccsé teszik. Mivel a könyvtárakban az ellenőrzött szótárak, tárgyszórendszerek rendszerint a nagyobb intézményekhez köthetők, ez a szerep lényegében a könyvtári rendszerek zászlóshajóinak (nemzeti könyvtárak, országos szakkönyvtárak, nagyobb közkönyvtárak) jelent feladatot.

A könyvtár akkor válhat a globális tudáscsere központi szereplőjévé, ha az általa közzétett stabil URI-kat mások fel is használják. A könyvtárnak tehát el kellene érni, hogy a szemantikus technológiák szabványosításával foglalkozó

szervezetek ne maguk kísérlejenek meg ilyen típusú tartalmat generálni, hanem a már meglévő tárgyszójegyzékek konvertálásával oldják meg az azonosítás problémáját. A szemantikus technológiákat fejlesztő szervezetek rendszerint a tudásreprezentáció technológiai kérdéseivel foglalkoznak, és alig vesznek tudomást a metaadat-tartalom hiányának problémájáról.

A globális tudáscsere igénye történeti korokon át folyamatosan fennállt. A kezdeti centralista utópiák után kiderült, hogy a decentralizált munka és a hálózatépítés az egyetlen útja, hogy ez megvalósuljon. A szemantikus technológiák lényegében erre épülnek, hiszen szabványok közzétételétől remélik a tudás alkalmazások fölötti megragadását. A szemantikus technológiáknak van egy olyan sajátossága, hogy mivel fokozottan támaszkodnak tudásszervezési rendszerekre, mindig is igényelni fognak professzionális hátteret. Minél inkább a használókra akarjuk hagyni a tartalmak és a metaadatok generálását, annál nagyobb szükség van a háttérben a kiterjedt, stabil és megbízható tudásszervezési rendszerekre. A szemantikus technológiák akkor „vihetők le” a használók szintjére, illetve akkor érhető el azok kiterjedt használata, ha a valóság bonyolult struktúráinak a tudásreprezentáció szintjén való formalizálását megcsináljuk helyettük. A FOAF erre is például szolgál. Mindez újabb okot ad arra, hogy a könyvtárak a meglévő tudásszervezési rendszereiket konvertálják a szemantikus technológiák igényeinek megfelelően.

Egyelőre nem látszik az esély, hogy a szemantikus web víziója – elvben és kicsinyített modellekben működő infrastruktúrája – valaha valósággá váljék-e. Mindenesetre, ha a világ ezt a lehetőséget hagyja elúszni, abban a könyvtáraknak és a könyvtárosoknak is igen nagy felelőssége lesz, mivel a szemantikusweb-technológiák többnyire könyvtáros szakértelmet igényelnek. Ha a szemantikusweb-technológiák – folytatva az eddigi tendenciákat – tovább terjednek, akkor viszont a digitális könyvtárak megőrizhetik, sőt növelhetik jelentőségüket a 21. században is. (Parker, Nitse, Flowers 2005)

A szemantikus világháló és vele a globális tudás-

csere nagy részben azon múlik, hogy mennyire terjednek el a szemantikát megragadni képes technológiák. A könyvtári hálózat, amellet, hogy a tudásreprezentáció tekintetében jelentős hagyományokkal és bőséges tapasztalattal rendelkező intézményrendszer, kiterjedt kapcsolati hálóval rendelkezik, amit mozgósíthat a szemantikusweb-szabványok minél kiterjedtebb használatát elősegítendő. A könyvtárak kapcsolati hálójának az adja az igazi erejét, hogy a leginkább tudásintenzív vállalatok és intézmények jelentik a partneri kört (egyetemek, kutatóintézetek, kormányzati szervek, nagyvállalatok, önkormányzatok stb.). A könyvtáraknak a közvetlen partnereiket kellene segíteni és támogatni abban, hogy képesek legyenek az általuk birtokolt tartalmakat a szemantikus technológiákkal kompatibilis módon közzétenni. Az ilyen típusú segítségnyújtás legjellemzőbb példája lehet a tudásszervezési rendszerek kidolgozása.

Mindehhez el kell érn, hogy a könyvtárat magát is partnernek tekintsék a fent felsorolt tudásintenzív szervezetek. Ehhez a könyvtárnak tudatosítania kell, hogy a tudás és az információ kezelésében jártas, magas szinten képzett munkatársakkal rendelkezik, így képes a vállalatoknál, különböző szervezeteknél felmerülő tudásreprezentációs problémák megoldására. A könyvtár megjelenésének a professzionalizmust és a legújabb technológiában való jártasságot kell sugallnia. Egyes vállalati könyvtárak esetében – részben a túlélésért folytatott küzdelem eredményeként – sikerült elérni, hogy a könyvtár a szervezetben megjelenő tudás szervezésének kulcsintézménye lett (Sajó 2003; Kóródy 2009). Más könyvtártípusok, és különösen a kisteleplési könyvtárak ugyanakkor jelenlegi állapotukban nagyon távol állnak attól, hogy velük a használói közösség a professzionális tudásszervező szerepet azonosítsa. (Péterfi, Tóth, Vidra 2008; Tóth 2009)

A szemantikus technológiák adaptálásához a hazai kis- és középvállalkozások jelenleg nem rendelkeznek megfelelő kapacitással és ráadásul mindez nem is jelent számukra prioritást. A könyvtárak ezt a potenciális használói kört is

megcélozhatnák olyan szolgáltatásokkal, amelyek a tudás vállalati szintű reprezentációjához jelentenek segítséget. A kis- és középvállalkozások jelenleg a könyvtárak számára más tekintetben is kiaknázatlan piacot jelentenek (Kiszl 2005).

Az eddig felvázolt jövőkép azt erősíti bennünk, hogy az új paradigma megjelenése a könyvtár-automatizálás időszakában elindult tendenciák erősödését hozza. (Tóth 2007) A könyvtárnak tudatosan fel kell vállalnia az információ és a tudás szervezésének az aktív, de periférikusabbnak tűnő szerepét. Periférikusság alatt értem egyrészt a „láthatatlanságot”, másrészt a „szemantika-ileg kódolt információ kizárólagos birtokosa” szerepről való tudatos lemondást. Mindkettő aspektus magyarázatra szorul.

A jelenlegi világhálóval való interakció során is jellemző, hogy a használó nem vesz tudomást arról, hogy mely tartalmat melyik szervezet, intézmény, vállalat vagy magánszemély tette számára hozzáférhetővé, és szeret úgy tekinteni a webre, mint egy hatalmas középpont nélküli szövegre, amelynek a lekereséséhez a nagy keresőrendszerek jelentik a belépési pontot. A szemantikus web világában az ember-gép interakció során a találati halmaz ideális esetben nem dokumentumok listája lesz, hanem konkrét információké, amelyek a legkülönbözőbb szolgáltatók által közzétett tartalmak alapján kerülnek prezentálásra.

A könyvtárak számára sokkoló lesz lemondani arról a pozícióról, amit a szemantikailag kódolt információ szinte kizárólagos birtoklása jelent. A könyvtárak tudásszervezési rendszereik közzétételével a legnagyobb hozzáadott értékkel bíró tartalmaikat teszik közkinccsé, és egyben lemondanak annak birtoklásáról. A metaadat-tartalmak konverziója többek között ezért sem megy zökkenőmentesen. (Coyle 2008; Burke 2009) Amennyiben a könyvtárak valóban elkötelezettek az információ szabad hozzáférhetővé tétele mellett, akkor ezt a lépést a használók érdekében meg kell tenniük. (Ráadásul a használók ezért nem lesznek hálásak nekünk, hiszen várhatóan nem is szereznek róla tudomást, hogy

a könyvtárnak köszönhetik a háttérben futó tudásszervezési rendszereket.)

Ahhoz, hogy a könyvtárak felvállalják, hogy megoldják a szemantikus technológiák számára a tudásszervezési rendszerek biztosításának és a fogalmak egyértelmű azonosításának problémáját, a jelenleginél intenzívebb együttműködést kell kialakítani a szemantikus technológiák szabványosításával foglalkozó szervezetekkel (Greenberg 2007). A könyvtáraknak nagyobb számban kellene magukat képviseltetni a szabványosítási testületek munkájában, hiszen ezek a jelenleg periférikusnak tűnő kérdések a szakma jövőjét határozhatják meg, amennyiben a szemantikus web lesz az új könyvtári világrend kialakulásának motorja. (Vö. Joint 2008) A szemantikusweb-fejlesztésekben való aktív részvétel nem csak a könyvtárnak mint intézményrendszernek, de a könyvtáros egyéni karrierjére nézve is megtermékenyítő erővel bír, hiszen olyan kihívásokat jelent, amelyben a tudásszervezéssel kapcsolatos ismeretek teljes skáláját fel kell vonultatni.

Jelenleg a könyvtárosok napi rutinszerű munkája szempontjából elenyésző jelentősége van a szemantikus technológiáknak. (Joint 2008) Ameddig a könyvtárak nem vállalnak intenzívebb szerepet a szemantikailag kódolt információk közzétételében, addig ebben a tekintetben nem is várható változás. Ugyanakkor a jövő informatikus-könyvtárosainak képzésében a jelenleginél sokkal nagyobb szerepet kell szánni a szemantikus technológiák ismeretének és gyakorlati alkalmazásának. Azt is be kell látni, hogy a szemantikus technológiák alkalmazásához elengedhetetlen „professzionális informá-

cióközvetítő” illetve „tudásmérnök” imázs csak a képzésen keresztül érhető el.

Azon túl, hogy a fent megfogalmazott lehetőségekkel a könyvtárnak élnie kell, nagyon sok – a szemantikus technológiák által generált – kihívással kell a könyvtáraknak és a technológiák fejlesztőinek szembenéznük. Kihívás magának a metaadat-infrastruktúrának az elkészítése és fenntartása, amely rendkívül idő- és erőforrásigényes. (Burke 2009) A létező tudásszervezési rendszerek és a metaadatok konverziója csak részben oldja meg ezt a problémát, ugyanis ahol maga a szemantika semmilyen formában nem került kódolásra, ott ezt meg kell tenni (és a megbízhatóság érdekében nem heurisztikák alkalmazásával, hanem manuálisan). A szemantikus web és a különböző forrásokból összekapcsolt adatok nagyon komoly szinten felvetik az adatbiztonság kérdését. A szemantikus web víziója szerint olyan információk is elérhetők lesznek, amelyeket közvetlenül nem adtak meg, csak más tartalmak alapján kikövetkeztethető. Ez mind a magánszemélyek, mind a cégek, mind pedig hivatalok, szervezetek, intézmények esetében nagyon komoly veszélyt jelent. A probléma megoldása részben technológiai kérdés. Nagyon sok olyan kihívás van, amelynek megoldása nem elsősorban a könyvtárak feladata, viszont a közreműködésére szükség van, amennyiben az intelligens világhálót szeretné a saját igényeihez és lehetőségeihez formálni. Szintén kihívást jelentenek olyan, korábban már részletesen bemutatott problémák, mint az ontológiai automatikus generálása, vagy a fogalmak egyértelmű azonosítása.

Köszönetnyilvánítás

A dolgozat megírásához nyújtott szakmai segítségért és a kritikai megjegyzésekért köszönetet mondok tanárainknak, kollégáimnak: Prof. Dr. Sebestyén Györgynek, Dr. Horváth Péternek, Dr. Fodor Jánosnak (ELTE); Dr. Sonnevend Péternek (Kaposvári Egyetem); Ládi Lászlónak (Könyvtári Intézet) Ragnar Audunsonnak, Nils Pharonak (Høgskolen i Oslo); Steve Peppernek (Ontopedia) és Dr. Kovács Zsigmondnak (Kovex Computer Kft.). Köszönetet mondok továbbá hallgatóimnak a Pécsi Tudományegyetem Felnőttképzési és Emberi Erőforrás Fejlesztési Karán. A kutatást ösztöndíjjal támogatta a Magyar Ösztöndíj Bizottság és a Norvég Kutatási Tanács (Norges Forskningsrådet).

Jegyzetek

- 1 Valamennyi W3C ajánlás elérhető magyarul a <http://www.w3c.hu/forditasok/> címen (2010. július 6.).
- 2 A DARPA a *Defence Advanced Research Project Agency* rövidítése, amely 2006-ig az amerikai védelmi minisztérium központi kutatásfejlesztési szervezete volt.

Irodalom

- ANBARASAN, E. (2000): Tim Berners-Lee: The Web's Brainchild. = Unesco Courier, September, 2000. http://www.unesco.org/courier/2000_09/uk/dires.htm [letöltve: 2010. január 11.]
- ALESSO, H. P. – SMITH, C. F. (2009): Thinking of the Web: Berners-Lee, Gödel and Turing. New Jersey : John Wiley & Sons, 289 p.
- BAKER, D. (2006): Digital library futures: a UK HE and FE perspective. = Interlending & Document Supply, 34. vol. 2006. 1. no. 4–8. p.
- BERNERS-LEE, T. (1998): Semantic Web Road Map, September, 1998. Draft. <http://www.w3.org/DesignIssues/Semantic.html> [letöltve: 2010. január 20.]
- BERNERS-LEE, T. – FIELDING, R. – IRVINE, U.C. – MASINTER, L. (1998): RFC2396. Uniform Resource Identifiers (URI): generic syntax, August, 1998. <http://www.ietf.org/rfc/rfc2396.txt> [letöltve: 2010. január 20.]
- BERNERS-LEE, T. – HANDLER, J. – LASSILA, O. (2001): The Semantic Web. = Scientific American, May 17. <http://www.scientificamerican.com/article.cfm?id=the-semantic-web> [letöltve: 2010. január 20.]
- BÍRÓ Sz. (2004): A szövegfeldolgozás modern eszközei – az SGML és XML nyelvek. = Tudományos és Műszaki Tájékoztatás, 51. évf. 2004. 10. sz. 453–459. p. http://tmt.omikk.bme.hu/show_news.html?id=3733&issue_id=455 [letöltve: 2010. január 20.]
- BÍRÓ Sz. (2005): Szövegfeldolgozás XML alapokon. Bp. : Neumann Kht., 208 p.
- BOLEY, H. et al. (2009): RIF Core Dialect W3C candidate recommendation, 1 October 2009. <http://www.w3.org/TR/2009/CR-rif-core-20091001/> [letöltve: 2010. január 4.]
- BRAY, T. – HOLLANDER, D. – LAYMAN, A. – TOBIN, R. – THOMPSON, H. S. (2009): Namespaces in XML 1.0 (Third ed.) W3C recommendation, 8 December, 2009. <http://www.w3.org/TR/2009/REC-xml-names-20091208/> [letöltve: 2010. január 4.]
- BRAY, T. – PAOLI, J. – SPERBERG–McQUEEN, C. M. (1998): Extensible Markup Language W3C recommendation, 10 February 1998. <http://www.w3.org/TR/1998/REC-xml-19980210> [letöltve: 2010. január 21.]
- BRAY, T. – PAOLI, J. – SPERBERG–McQUEEN, C. M. – MALER, E. – YERGEAU, F. (2008): Extensible Markup Language (XML) 1.0 (Fifth Edition) W3C recommendation, 26 November, 2008. <http://www.w3.org/TR/2008/REC-xml-20081126/> [letöltve: 2010. január 21.]
- BRICKLEY, D. – GUHA, R. V. (2004): RDF Vocabulary Description Language 1.0: RDF schema W3C recommendation, 10 February, 2004. <http://www.w3.org/TR/rdf-schema/> [letöltve: 2010. január 21.]
- BRICKLEY, D. – MILLER, L. (2010): FOAF vocabulary specification 0.97. Namespace document, 1 January, 2010 – 3D Edition <http://xmlns.com/foaf/spec/> [letöltve: 2010. január 21.]
- BURKE, M. (2009): The semantic web and the digital library. = Aslib Proceedings: New Information Perspectives, 61. vol. 2009. 3. no. 316–322. p.
- CAMPBELL, D. G. – FAST, C. V. (2004): Academic libraries and the semantic web: what the future may hold for research-supporting library catalogues. = Journal of Academic Librarianship, 30. vol. 2004. 5. no. 382–390. p.
- CANTARA, L. (2006): Encoding controlled vocabularies for the Semantic Web using SKOS Core. = OCLC Systems & Services, 22. vol. 2006. 2. no. 111–114. p.
- CONOLLY, D. et al. (2001): DAML+OIL (March 2001) reference description. W3C Note, 18 December, 2001. <http://www.w3.org/TR/daml+oil-reference> [letöltve: 2010. január 21.]
- COYLE, K. (2008): Meaning, technology, and the Semantic Web = Journal of Academic Librarianship, 34. vol. 2008. 3. no. 263–264. p.
- CRAWFORD, W. – GORMAN, M. (1995): Future Libraries: Dreams, Madness and Reality, Chicago and London, American Library Association, 198 p.
- DACONTA, M. C. – OBRST, L. J. – SMITH, K. T. (2003) : The Semantic Web: A Guide to the Future of XML, Web Services and Knowledge Management, Indianapolis, Wiley, 265 p.
- DODDS, L. (2005): Introducing SPARQL: querying the semantic web. 2005. November <http://www.xml.com/pub/a/2005/11/16/introducing-sparql-querying-semantic-web-tutorial.html> [letöltve: 2010. január 21.]
- FERRAN, N. – MOR, E. – MINGUILLON, J. (2005): Towards

- personalization in digital libraries through ontologies. = *Library Management*, 26. vol. 2005. 4–5. no. 206–217. p.
- FÜLÖP Cs. – KOVÁCS L. – MICSIK A. (2004): A metaadatsémák és a szemantikus web: egységesítés és specializáció a metaadatok világában. = *Tudományos és Műszaki Tájékoztatás*, 51. évf. 2004. 7. sz. 276–284. p. http://tmt.omikk.bme.hu/show_news.html?id=3648&issue_id=452 [letöltve: 2010. január 21.]
- GILCHRIST, A. (2003): Thesauri, taxonomies and ontologies: an etymological note. = *Journal of Documentation*, 59. vol. 2003. 1. no. 7–18. p.
- GOOSSENS, P. (2003): ELAG 2002: report of a library systems seminar on the Semantic Web. = *Program: electronic library and information systems*. 37. vol. 2003. 4. no. 251–253. p.
- GOTTDANK, T. (2005): Szemantikus web. Bevezetés a tudásalapú Internet világába. Bp. : Computerbooks, 282 p.
- GREENBERG, J. (2007): Advancing the Semantic Web via library functions. = *Cataloging & Classification Quarterly*, 43. vol. 2007. 3–4. no. 203–225. p.
- GREENBERG, J. – MÉNDEZ, E. szerk. (2007): *Knitting the Semantic Web*, Binghamton, New York, Haworth Information Press, 257 p.
- HAN, Y. (2005): A RDF-based digital library system. = *Library Hi Tech*, 24. vol. 2005. 2. no. 234–240. p.
- HECHT, R. – HASLHOFER, B. (2005): Joining the BRICKS Network – A piece of cake. = *The International Conference EVA 2005*. Moscow. http://www.cs.univie.ac.at/upload/550/papers/Joining_the_BRICKS_Network_-_A_Piece_of_Cake.pdf [letöltve: 2010. január 22.]
- HERMAN I. (2006): Szemantikus web: egy rövid bevezetés. Elhangzott az I. Magyar Webkonferencián 2006. március 18-án. <http://www.w3.org/2006/Talks/0318-Budapest-IH/cikk.html> [letöltve: 2010. január 22.]
- HORVÁTH T. – SÜTHEŐ P. (2001): A tartalmi feltárás. = *Könyvtárosok kézikönyve 2. Feltárás és visszakeresés*. Szerk. Horváth T. ; Papp I. Bp. : Osiris, 2001. 35–186. p.
- HORVÁTH Z. (2006): Taxonómia – az egyezményes nyelvek szerepe és rokonságai – útközben a szemantikus webhez. = *Tudományos és Műszaki Tájékoztatás*, 53. évf. 2006. 5. sz. 211–223. p. http://tmt.omikk.bme.hu/show_news.html?id=4410&issue_id=472 [letöltve: 2010. január 22.]
- ISAAC, A. – PHIPPS, J. – RUBIN, D. (2007): SKOS use cases and requirements. W3C working draft, 16 May 2007. <http://www.w3.org/TR/2007/WD-skos-ucr-20070516/> [letöltve: 2010. január 22.]
- ISAAC, A. – SUMMERS, E. (2009): SKOS Simple Knowledge Organization System Primer W3C working draft, 15 June, 2009. <http://www.w3.org/TR/skos-primer/> [letöltve: 2010. január 22.]
- JOINT, N. (2008): The practitioner librarian and the semantic web: ANTAEUS = *Library Review*, 57. vol. 2008. 3. no. 178–186. p.
- KISZL P. (2005): *Üzleti információ, céginformáció és a könyvtárak*, Budapest, Traduirex kiadó, 235 p.
- KLYNE, G. – CARROLL, J. J. (2004): Resource Description Framework (RDF) concepts and abstract syntax. W3C recommendation, Febr. 10, 2004. <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/> [letöltve: 2010. január 22.]
- KÓRÓDY J. (2009): Gyakorlati marketing vállalati könyvtárakban = *Tudományos és Műszaki Tájékoztatás*. (56. köt.) 1. sz. 18-26. p. http://tmt.omikk.bme.hu/show_news.html?id=5042&issue_id=500 [letöltve: 2010. január 19.]
- KRAUSE, J. (2008): Semantic heterogeneity: comparing new semantic web approaches with those of digital libraries. = *Library Review*, 57. vol. 2008. 3. no. 235–248. p.
- KRUK, S. R. (2008a): Extensible Access Control (EAC) ontology specification. Specification document, 01 July, 2008. <http://www.jeromedl.org/eac/1.0/spec/index.html> [letöltve: 2010. január 22.]
- KRUK, S. R. (2008b): JeromeDL ontology specification. Specification document. 02 April, 2008. <http://www.jeromedl.org/ontology/2.0/spec/index.html> [letöltve: 2010. január 22.]
- KRUK, S. R. – CZYGAN, M. – GZELLA, A. (2008): JeromeDL: semantic and social technologies for improving user experience in digital libraries. [Galway] : DERI – Digital Enterprise Research Institute. <http://library.deri.ie/resource/rsYUKcOR> [letöltve: 2010. január 22.]
- KRUK, S. R. – KRUK, E. – STANKIEWICZ, K. (2008): Results of the evaluation of the semantic and social technologies for digital libraries. [Galway] : DERI – Digital Enterprise Research Institute. <http://library.deri.ie/resource/ARfuVUI8> [letöltve: 2010. január 22.]
- LU, S. – Dong, M. – FOTOUCHI, F. (2002): The semantic web: opportunities and challenges for the next-generation Web applications. = *Information Research*, 7. vol. 2001. 4. no. <http://informationr.net/ir/7-4/paper134.html> [letöltve: 2010. január 23.]
- LYTRAS, M. – SICILIA, M. A. – DAVIES, J. – KASHYAP, V. (2005): Digital libraries in the knowledge era. *Knowledge management and Semantic Web technologies* = *Library Management*, 26. köt. 2005. 4–5. no. 170–175. p.

- MACGREGOR, G. (2008): Introduction to a special issue on digital libraries and the semantic web: context, applications and research. = *Library Review*, 57. vol. 2008. 3. no. 173–177. p.
- MALMSTEN, M. (2008): Making a library catalogue part of the semantic web. = *Proceedings of international conference on Dublin Core and metadata applications*. Berlin, 22–26 September, 2008. <http://www.kb.se/dokument/Libris/artiklar/Project%20report-final.pdf> [letöltve: 2010. január 4.]
- MANOLA, F. – MILLER, E. (2004): RDF Primer. W3C recommendation, 10 February, 2004. <http://www.w3.org/TR/rdf-primer/> [letöltve: 2010. január 23.]
- McCATHIE, N. – MÉNDEZ, E. (2007): Library cards for the 21st century = *Cataloging & Classification Quarterly*, 43. vol. 2007. 3–4. no. 21–45. p.
- McGUINNESS, D. L. – Van HARMELEN, F. (2004): OWL Web Ontology Language – Overview, W3C recommendation, 10 February, 2004. <http://www.w3.org/TR/owl-features/> [letöltve: 2010. január 23.]
- MILES, A. – BECHHOFFER, S. (2009b): SKOS Simple Knowledge Organization System Reference, W3C Recommendation, 18 Aug, 2009 <http://www.w3.org/TR/2009/REC-skos-reference-20090818/> [letöltve: 2010. január 4.]
- MOLDOVÁN I. (1999): A számítógépen tárolt dokumentumok időtálló formátuma, az SGML/XML. Egyszeres felvitel, sokirányú felhasználás. Bp. 1999. május 18. MKE Elektronikus Könyvtári Szekció. http://www.bibl.u-szeged.hu/mke_eksz/rendezvenyek/SGML99/index.html [letöltve: 2010. január 23.]
- MSZ 3418 (1987): MSZ 3418-87 Magyar nyelvű információkereső teauruszok szerkezete, részei és formái. 1988. január 1. Bp. : Magyar Szabványügyi Hivatal. [1987], 10. p.
- NUCCI, M. – HAHN, D. – BARBERA, M. (2008): The Talia library platform. Rapidly building a digital library on Rails. = 4th Workshop on Scripting for the Semantic Web colocated with ESWC 2008 June 2, 2008 Tenerife, Spain. <http://www.semanticscripting.org/SFSW2008/papers/4.pdf> [letöltve: 2010. január 23.]
- NUCCI, M. – BARBERA, M. – MORBIDONI, C. – HAHN, D. (2008): A Semantic Web Powered Distributed Digital Library System. = *Open Scholarship: Authority Community and Sustainability in the Age of Web 2.0. Proceedings ELPUB 2008 Conference on Electronic Publishing – Toronto, Canada – June, 2008*. 130–139. p. http://elpub.scix.net/data/works/att/130_elpub2008.content.pdf [letöltve: 2010. január 23.]
- PARKER, K. R. – NITSE, P. S. – FLOWERS, K. A. (2005): Libraries as knowledge management centers = *Library Management* (26. vol.) 4-5. no. 176-189. p.
- PASSIN, T. B. (2004): *Explorer's Guide to the Semantic Web*. Greenwich : Manning, 281 p.
- PATAKI E. ford. (2004a): Az RDF Erőforrás Leíró Keretrendszer alapfogalmai és absztrakt szintaxisa. W3C ajánlás, 2004-02-10. <http://www.w3c.hu/forditasok/RDF/REC-rdf-concepts-20040210.html> [letöltve: 2010. január 23.]
- PATAKI E. ford. (2004b): OWL Web Ontológia Nyelv – Áttekintés, 2004-09-19. <http://www.w3c.hu/forditasok/OWL/REC-owl-features-20040210.html> [letöltve: 2010. január 23.]
- PATAKI M. (2005): W3C ajánlások magyarul. = *Tudományos és Műszaki Tájékoztatás*, 52. évf. 2005. 9. sz. 430. p. http://tmt.omikk.bme.hu/show_news.html?id=4051&issue_id=465 [letöltve: 2010. január 23.]
- PATAKI M. – RICHTER V. (2008): A W3C szabványosítási törekvései. = *Networkshop 2008. Dunaújvárosi Főiskola 2008. március 17-19.* <https://nws.niif.hu/hcd2008/docs/phu/096.pdf> [letöltve: 2010. január 23.]
- PEPPER, S. (2008a): Topic Maps and the Semantic Web. = *Topic Maps and All That*, 2008-05-11 <http://topicmaps.wordpress.com/2008/05/11/topic-maps-and-the-semantic-web/> [letöltve: 2010. január 23.]
- PÉTERFI R. – TÓTH M. – VIDRA Sz. F. (2008): Kisvárosi könyvtárak szerepvállalásai = *Könyvtári Figyelő*, 54. köt. 2008. 1. sz. 35–54. p. http://www.ki.oszk.hu/kf/e107_plugins/content/content.php?content.104 [letöltve: 2010. január 19.]
- PRUD'HOMMEAUX, E. – SEABORNE, A. (2008): SPARQL Query Language for RDF. W3C Recommendation, 15 January, 2008. <http://www.w3.org/TR/rdf-sparql-query/> [letöltve: 2010. január 15.]
- SAJÓ A. (2003): Információ- és tudásmenedzsment elemek integrálása a könyvtári munkába. Hogyan váljunk vállalati könyvtárból információ- és tudásközponttá? = *Tudományos és Műszaki Tájékoztatás*, 50. köt. 2003. 12. sz. 491–506. p. http://tmt.omikk.bme.hu/show_news.html?id=3452&issue_id=446 [letöltve: 2010. január 19.]
- SALGÁNÉ MEDVECZKI M.: (2004): Az XML: új perspektívák a könyvtár-informatikában. = *Tudományos és Műszaki Tájékoztatás*, 51. évf. 2004. 2. sz. 61–71. p. http://tmt.omikk.bme.hu/show_news.html?id=3541&issue_id=448 [letöltve: 2010. január 23.]
- SALGÁNÉ MEDVECZKI M.: (2005): Our electronic era and bibliographic informations computer-related bibliographic data formats, metadata formats and BDML. = *Librarianship*

- in the information age. Proceeding of the 13th International BOBCATSSS Symposium, Bp. : ELTE, 2005. 211–231. p.
- SURE, Y. – STUDER, R. (2005): Semantic Web technologies for digital libraries = Library Management (26. köt.) 4-5. sz. 190-195. p.
- SZAKADÁT I. (2006): MEO Magyar Egységes Ontológia szakmai zárójelentés. Bp. 2006-12-18. 13 p. http://ontologia.hu/meo/docs/meo_report_final_2006 [letöltve: 2010. január 23.]
- SZAKADÁT I. (2007a): Egyben az egész egytől egyig. Bp. : Typotex, 2007. 142. p.
- SZAKADÁT I. (2007b): A nép szavai Címkék és könyvtárak – a szemantikus web ígérete és valósága = A Web 2.0 jövője és a szemantikus web. A 2007. április 27-i in4 konferencia tanulmányai. Budapest, Bibliopolisz, 2007. 29–55. p.
- SZEREDI P. – LUKÁCSY G. – BENKŐ T. (2005): A szemantikus világháló elmélete és gyakorlata. Bp. : Typotex, 2005. 501 p.
- TÓTH K. (2003): XML alkalmazások a HunTéka könyvtári rendszerben. Bp. : MTA SZTAKI. 2003. <http://aromo.aszi.sztaki.hu/~tothk/Hunteka/XMLHunteka.htm> [letöltve: 2010. január 23.]
- TÓTH M. (2007): A szemantikus web új lehetőségek és kihívások könyvtári környezetben = AGRIA MEDIA 2006. (Konferenciakötet) Eger, 2007. 151–157. p.
- TÓTH M. (2009): Pushing the boundaries of accessibility governmental and professional efforts on ensuring equal access to information to rural library users (1997-2007) = NETCOM – Networks and Communication Studies. 23. köt. 1–2. sz. Hungarian information society – Selected studies. 85–110. p.
- UNGVÁRY R. (2004): Tezaurusz és ontológia, avagy a fogalmi ismertetőjegyek generikus öröklődésének formalizálása. = Tudományos és Műszaki Tájékoztatás, 51.évf. 2004. 5. sz. 175–191. p. http://tmt.omikk.bme.hu/show_news.html?id=3615&issue_id=450 [letöltve: 2010. január 23.]
- VAS B. (2005): A hagyományos tartalomfeltárástól a szemantikus web alkalmazásáig. Szakdolgozat. Szegedi Tudományegyetem Könyvtártudományi tanszék, 52 p. [Kézirat]
- VICKERY, B. C. (1997): Ontologies. = Journal of Information Science, 23. vol. 1997. 4. no. 277–286. p.
- VÖLKEI, M. – KRÖTZSCH, M. – VRANDECIC, D. – HALLER, H. – STUDER, R. (2006): Semantic Wikipedia. = www2006 proceedings. The 15th International World Wide Web Conference, Edinburgh, Scotland <http://www2006.org/programme/files/pdf/4039.pdf> [letöltve: 2010. január 23.]
- WARREN, P. – ALSMEYER, D. (2005): Applying semantic technology to a digital library: a case study = Library Management, 26. köt. 2005. 4–5. sz. 196–205. p.
- WORONIECKI, T. – KRUK, S. R. (2007): JeromeDL Ontology Specification Specification Document, 10 March 2007. http://www.jeromedl.org/structure/spec/index.html#term_previousVersion [letöltve: 2010. január 23.]
- YU. L. (2007): Introduction to the Semantic Web and Semantic Web Services. Chapman & Hall/CRC, 333 p.

Régi könyvek leírási szabványa

Elkészült a régi könyvek bibliográfiai leírásának végleges magyar nyelvű változata. A szabályzat alapja az ISBD(A) 2006-os kiadása, de szerkezetében a magyar szabályzatokhoz illeszkedik, és számos, a régi nyomtatványokkal kapcsolatos fogalom meghatározásával, valamint az alkalmazást megkönnyítő példával egészült ki. A szabályzatot kidolgozó bizottság a közelmúltban nyilvános vitára bocsátotta a tervezetet, és a végleges szöveg kialakításánál figyelembe vette a beérkezett hozzászólásokat.

A végleges szöveg leölthető a Könyvtári Intézet honlapjáról:
<http://www.ki.oszk.hu/107/download.php?list3>.

Közvetlen elérhetősége:

<http://www.ki.oszk.hu/107/dowbload.php?view.573>

(Katalist, 2010. júl. 14. Keveházi Katalin tájékoztatása alapján)