Szemantikus web és könyvtár: egy konferenciasorozat margójára

(Hat előadás a SWIB 2017/2018-as konferenciáiról)

A kilencvenes években − az internet elterjedésével − az a kép alakult ki, hogy a könyvtáraknak meg kell mutatkozniuk a hálózaton és igyekezniük kell mielőbb elérhetővé tenni honlapjaikat és persze katalógusukat a világhálón. Akkor még nem látszódott, hogy a könyvtárak jövője nem csupán a hálózaton, hanem egyenesen a* hálózatban *van. Ez nem üres szillogizmus, a kettő közt óriási minőségi különbség van. Jó néhány éve már, hogy könyvtáros berkekben felbukkant, a jövő egyik legjelentősebb ránk váró feladata a webtér integrációja lehet. Hiszen a „hálózatba vetettség" további körülménye, hogy átalakul magának a katalógusnak a fogalma is. A könyvtárak és állományaik, leíró adataik egyként az óriási hipertext-tér részei lesznek, mégpedig olyan módon, hogy eddigi „kétdimenziós" jelenlétük „háromdimenzióssá" válik.

Tárgyszavak: *szemantikus web, linked data, SWIB-konferenciák*

A könyvtári katalogizálás újfajta szemlélete lehetővé teszi, hogy az egyes könyvtárakban leírt objektumok összekapcsolhatók legyenek, mégpedig szemantikus módon az egész hálózati hipertexttérrel és ilyen módon a könyvtári adatfeltárás maga is részévé válik ennek a webtérnek, mintegy a webtérbe beágyazva (a „linked data" technológia könyvtári vetülete). Ennek a gondolkodásnak világos folyománya az új koncepció megjelenése, az FRBR (Functional Requirements for Bibliographic Records), amely a bibliográfiai tételek funkcionális követelményeire vonatkozó model, s ami elméleti alapja lett az ISBD-t majd felváltó RDA (Resource Description and Access) szabványnak és a várhatóan a MARC helyébe lépő BIBFRAME-nek.

Ma már jól látszódik, hogy elindult a világ ezen az úton, s azt is észrevehetjük, hogy a „felhőbe" költöző könyvtári rendszerek és a területi, szakirányú vagy éppen nemzeti felhő alapú platformok kialakulása minderre még ösztönzőleg is hatott. Úgy véljük, nem lesz ez másképp nálunk sem, a hazai szakemberek munkálkodása nyomán is egyre másra születnek a híradások, előadások és dolgozatok (és talán projektek is) ebben a témában.[1]

Így a könyvtári területen is megjelentek a konkrét szemantikus webes fejlesztések. Mindezekről az egyik legfontosabb hely a mai Európában, ahol „élőben" tájékozódhatunk, az ún. SWIB konferenciasorozat. Az eredetileg német, majd gyorsan nemzetközivé váló *„Semantic Web in Bibliotheken"* (ill. *Semantic Web in Libraries*) konferenciasorozat évek óta szolgálja ezt az ügyet. A 2009-ben elindult konferenciák programja nagyobb részt követhető volt a weben is, illetve az előadások jó része megtalálható a világhálón is, beleértve a 2017-es és a legutóbbi 2018-as rendezvényt is.[2]

Ez az írás tulajdonképpen a legutóbbi két konferenciáról emel ki hat előadást, hogy konkrétan is bemutassa őket, és egyben kedvet is csináljon a teljes előadások, vagy akár az egész 2017-es illetve 2018-as konferenciaanyag részletes tanulmányozásához.

ꕥ ꕥ ꕥ ꕥ ꕥ

Dr. Mia Ridge: Libraries & their communities: participation from town halls to mobile phones

A könyvtárak és felhasználóik: közreműködők a városházától a mobil eszköz-használókig
(Dr. Mia Ridge 2015-től a British Library's Digital Scholarship csoportjának digitális kurátora az ún. Western Heritage gyűjteményeiben)
https://www.youtube.com/watch?v=duiUgQS2tW4&feature=youtu.be

Mia Ridge bevezetőjében a könyvtári alaphelyzetet tárja elénk. Óriási a verseny a katalógusok és a közösségi közreműködésekre építő projektek között. A gyakorlati tapasztalat azt mutatja, hogy csak akkor tudnak a könyvtárosok lépést tartani a forradalmi fejlődésekkel, ha bevonják az olvasóikat is a feldolgozói munkafolyamatba és nyilván irányítják őket, biztosítják számukra a megfelelő programokat, kereteket.

A cél eléréséhez sokszor bátorítják és inspirálják a közreműködőket, mert sokan kishitűek a saját tudásukkal kapcsolatban, nem értik miért lehet jó a könyvtár számára az ő esetleges szaktudásuk. A könyvtáros feladata, hogy úgy fogalmazza meg a projekteket, hogy élvezhető legyen, sikerélményt adjon, és a mellett, hogy az olvasók kényelemben otthon fejlesztik a tudásukat, meg is oszthatják később publikáció formájában, vagy a könyvtári fórumokon. Meg kell érteni, hogy mi az önkéntes kollégák motivációja, mi okozhat számukra örömet, mi lehet vonzó úgy, hogy közben sokat tanulnak belőle.

Példákon keresztül bemutat több „mikro feladatot", hogyan lehet a mindennapi embert bevonni az adatbázis-építésbe akár egy telefon segítségével: családfakutatásnál ilyen lehet a nevek átírása kézírásból szkennelt képek alapján, vagy a Wikipédiában nevek egységesítése, kikeresése. Rosszul OCR-ezett szövegek kitisztázása, vagy képadatbázisban a címkézés, tárgyszavazás.

Tavaly az Oxford English Dictionary átdolgozása kapcsán kérték a fiatal közösséget[3], hogy Twitteren írjanak nekik szavakat arról a vidékről, ahol élnek, ezzel is elősegítve a következő szótár összeállítását.

Végül a British Library projektjét az **„In the Spotlight"**[4] névre keresztelt közösségi programot mutatja be részletesebben. Itt kb. 230 ezer nyomtatott színlap várja az olvasókat, hogy átírják. Az önkéntes választhat, hogy csak a címet, a személyneveket vagy a műfajt írja be egyesével, de ezt megteheti egyszerre is, vagy laponként is. Eddig 8-900 regisztrált önkéntesük van és már 126 ezer beavatkozás, kiegészítés történt (ennyi szerkesztett címke.) Az olvasók kommentekben is tudnak kapcsolatot tartani, rengeteget segítenek a könyvtárosoknak és így végül sokkal több szempont szerint lesznek visszakereshetők a színi lapok.

Zárásként megfogalmazza a könyvtári közösségek kihívásait: nyitnunk kell a közösség felé, mert kevesen vagyunk a rengeteg adat feldolgozásához, és egyre nehezebbé válik kezelni a szemantikus webet. Viszont ki kell mondani azt is, hogy a mesterséges intelligencia alapú tudás sosem (?) helyettesítheti az emberi ellenőrzést. A rengeteg nyílt adathoz rengeteg emberre van szükség, akik feldolgozzák őket, úgy, hogy a következetesség és ellenőrzöttség elve ne csorbuljon!

ꕥ ꕥ ꕥ ꕥ ꕥ

Osma Suominen: Finnish National Bibliography Fennica as Linked Data

A Finn Nemzeti Bibliográfia mint Linked Data
(Osma Suominen informatikus szakember a Finn Nemzeti Könyvtárban, jelenleg automatizált tárgyi indexelés kialakításán és bibliográfiai adatok nyilvánossá tételén dolgozik)
https://www.youtube.com/watch?v=sLMxALIQKmQ&feature=youtu.be

A prezentáció a Finn Nemzeti Bibliográfia – FENNICA[5] – mint Linked data típusú project aktuális helyzetképét tárja az érdeklődők elé. Az előadó szinte missziójának tekinti a nemzeti könyvtár adatainak nyilvánossá[6] tételét, mely hatalmas mennyiségű információ korábban irreleváns volt a webhasználók körében. Ennek, és általánosságban a könyvtári tudás megosztásának fontosságát is hangsúlyozza, amikor a bibliográfiai adatok közzétételének okait sorolja fel.

Beszél a projekt indulásáról, amikor is adva volt több mint egy millió bibliográfiai rekord, 125 ezer személynév, 40 ezer intézményi név és 30 ezer tárgyszó authority rekordokban. Ahhoz, hogy ezek Linked dataként legyenek láthatók, meg kellett szüntetni az addigi adatszerkezetet, majd az elemeket újra kellett strukturálni. A MARC-rekordok konverziójának lépéseiről és az ehhez használt alkalmazásokról remek folyamatábrát mutat a hallgatóknak. A munkamenet egyes állomásairól részletesen is beszámol, megemlítve a járulékos feladatokat (mint külső szótárak hozzákapcsolása) vagy a felmerülő problémákat (például a duplikátumok kezelésének és kiszűrésének kérdése).

A prezentációban erről az általuk kialakított új adatmodellről jól átlátható, informatív ábrát is kapunk. Ebben központi elemeként a művet vették alapul, és ehhez gráfszerűen társították a számukra fontos kiegészítő információkat. Ezen információk forrásai egyrészt a saját katalógusukban szereplő MARC-rekordok voltak, ezekből nyerték ki

például a közreműködő személyek nevét (szerző, fordító vagy szerkesztő), a mű témáját leíró tárgyszavakat, kulcsszavakat, a mű náluk meglévő fizikai példányainak adatait stb. Másrészt források voltak egyéb adatbázisok, katalógusok, névterek (többek között a Library of Congress Subject Heading, a Finn Földrajzi névnyilvántartás vagy a Wikidata) is, az ezekben fellelhető ismereteket is hozzákapcsolták a művekhez. Az ábrán megjelennek továbbá olyan, már meglévő kiépítés alatt álló információforrások is, amelyekhez a közeljövőben szeretnék kiépíteni a kapcsolatokat

Az előadás végén röviden vázolja a távlati céljaikat, mint például az adatok tovább gazdagítása és letisztázása, még több hivatkozás beépítése más Linked data állományokra. Tervezik a VIOLA zenei gyűjtemény vagy az ARTO cikkadatbázis későbbi nyilvánossá tételét is a projekt keretein belül.

ꕥ ꕥ ꕥ ꕥ ꕥ

Osman Souminen: Annif: leveraging bibliographic metadata for automated subject indexing and classification

Annif: bibiográfiai metaadatok alkalmazásának előnyei az automatizált tartalom indexelésében és osztályozásában
(Osman Souminen Finn Nemzeti könyvtár informatikusa)
https://youtu.be/lSrFP3D-uTg?t=2

Osman Souminen a Finn Nemzeti Könyvtár munkatársaként két évvel ezelőtt kezdte el fejleszteni azt az Annif névre hallgató programot, amely arra hivatott, hogy meggyorsítsa és segítse az automatizált osztályozást és tartalomindexelést.

A konferencián elhangzott előadásának bevezetőjében röviden ismerteti, hogy honnan származik az Annif fejlesztésének ötlete[7]. A Finn Nemzeti Könyvtár 2013-ban létrehozta a finn könyvtárak, múzeumok közös keresőfelületét, a Finna-t[8], amely kb. 15 millió rekordot és ezek tárgyszavait tartalmazza. Az Annif prototípusa a Finna dokumentumainak metadatait használta fel, hogy teljes szövegeket indexeljen és osztályozzon.

Az előadásban bemutatta, hogyan épül fel ez a nyílt forráskódú eszköz, milyen fejlesztéseken esett át, milyen algoritmusokat alkalmaz és milyen teszteléseket hajtottak már végre vele.

Az Annif újonnan továbbfejlesztett változata három különböző algoritmust használ. Az algoritmusok közül kettő asszociatív, egy pedig lexikai megközelítésű. Az asszociatív közül az egyik a TF-IDF, amely a szövegben a kulcsszavak sűrűségét méri, nem számszerűen, hanem fontosság szerint állítja fel az eredményeket. Emellett használ még egy Fast Text nevű gépi tanulási modellt és lexikai megközelítéssel egy új-zélandi fejlesztésű MAUI nevű modellt. A program tervezésekor figyeltek arra, hogy többnyelvű legyen, egyelőre svéd, finn és angol nyelven elérhető és támogat SKOS-sémákban vagy egyszerű TVS-formátumban, és használható egy parancssori felületen vagy egy mikroszolgáltatás-stílusú REST API-n keresztül.

Beszél arról, hogyan végeztek egy összehasonlítást, ahol 50 különböző típusú dokumentum (cikkek, doktorik, könyvek) teljes szövegét indexálták négy különböző finn adatbázisból (Arto, JYU theses, Asklib, Satakunann kansa). Azt vizsgálták, hogy melyik algoritmussal érik el a legpontosabb eredményeket, tárgyszavakat. Az indexáláshoz a finn YSO-t (Általános Finn Ontológiát) használták. Megfigyelésük alapján egyéni algoritmusként a MAUI volt a legjobb, de összességében együtt érik el a leghatékonyabb elemzést.

Tesztelései között volt a finn Wikipédia kb. 410 000 szócikkének Annif általi automatikus indexelése, amely gépi módon kb. 7 órát vett igénybe, sokkal kevesebb időbefektetést igényelt, mintha ember készítette volna. Az eredményben pedig kijött, hogy melyek a leggyakrabban előforduló tárgykörök, amelyek előfordulnak a finn szócikkek között. Ezek között említi a focit, a jéghokit, a pop zenét és a hadihajókat.

Zárásként elmondta, hogy elérhető a program egy mobil applikációként, amely telefonra letöltve, OCR-t használva, szöveget elemez és tárgyszavaz, de elérhető mobil web appként[9], emellett még egy böngészőbe épített bővítménnyel Annif API elemzést követően könyvajánlat is kérhető a Finna.fi-ről. Az Annif dokumentációja pedig megtalálható a Github-on[10].

ꕥ ꕥ ꕥ ꕥ ꕥ

Dario Taraborelli: Unlocking Citations from tens of millions of scholarly Papers

Több tíz millió tudományos kiadvány hivatkozásainak szabadon hozzáférhetővé tétele

(Wikimedia Foundation, United States of America. 'Social computing' kutató, a tudomány szabadságának [open knowledge] szószólója San Franciscoból.)
https://www.youtube.com/watch?v=xcLv0oPCU0A&feature=youtu.be

Dario Taraborelli az igazgatója és kutatásvezetője a Wikimedia Alapítványnak, annak a non-profit szervezetnek, mely a Wikipédiát és testvérprojektjeit működteti. Az Initiative for Open Citations egyik alapító tagja, aki olyan rendszereket és programokat tervezett, amelyek felgyorsítják a tudományos eredményekhez való hozzájutást.

Több neves európai egyetemen (University College London, University of Surrey, Sciences Po, Paris Diderot University stb.) kutatói és tanári állást is betölt.

A hivatkozások az alapjai annak, hogy honnan tudjuk, amit tudunk. Ugyanakkor a tudományos életben a megbecsülést az idézések száma jelenti. Éppen ezért nem mindegy, hogy honnan, kitől ered egy idézet, mint ahogy az sem, hogy ezt hányan tartják értékesnek, s hivatkoznak rá maguk is. Ráadásul a tudományos eredményeket felmutató kutatások egy részét az adófizetők pénzéből finanszírozzák, ezért mindenki számára elérhetővé kellene tenni a kutatások eredményeit, s a rájuk való hivatkozások mérését.

Hogyan tudnánk ezeket az adatokat, az idézetekhez való hozzáférést nyilvánossá tenni, szerzői jogi megszorítások nélkül? Az itt összefoglalt előadás egy olyan kezdeményezést mutat be, mely ezt a problémát igyekszik megoldani.

Az Initiative for Open Citations (I4OC, https://i4oc.org) egy együttműködés, mely a nagy kiadók, kutatók, tudományos munkát végző szakemberek és egyéb érdekszervezetek között jött létre azzal a céllal, hogy támogassák a korlátozások nélküli hozzáférést a tudományos művek idézési adataihoz. A kezdeményezés kiindulópontja, hogy a legtöbb kiadó előre letétbe helyezi a hivatkozásainak adatait a Crossref-be (https://www.crossref.org/). A Crossref egy nonprofit szervezet, mely azért jött létre, hogy javítsa a tudományos kommunikációt. Jelenleg 103 millió folyóiratcikk, könyv, szabvány található benne, s évente kerülnek bele az újabb adatok. Az első lépés a megvalósításban az, hogy meg kell győzni néhány befolyásos kiadót, hogy adataikat „nyissák meg", tegyék hozzáférhetővé mindenki számára a Crossrefben. Ha ez megtörténik, akkor az I4OC adatbázis „le tudja aratni" ezeket az információkat, s ezt követően publikussá tudja tenni őket. A kezdeményezés indulásakor, 2016 szeptemberében a Crossrefben található hivatkozásoknak csak mindössze egy százaléka volt nyilvánosan elérhető. Hat hónappal a kezdeményezés indulása után több mint 40% lett ez az arány. Természetesen ezt az arányt folyamatosan növelni szeretnék, erre hivatott az I4OC, mely egy nagy szövetséget hozott létre kiadók, kutatók, szervezetek között.

Az előadó bemutatott még néhány hasonló kezdeményezést, például az Open Citation Corpust (http://opencitations.net/corpus), mely egy széles körű, nyilvános gyűjteménye a hivatkozásoknak. Beszélt még a VOSviewer-ről (http://www.vosviewer.com/), mely egy szoftver segítségével vizualizálja a tudományos térképet, az idézési hálózatot láthatóvá teszi. Adatai a Web of Science és a Scopus adatbázisokból származnak. Ezt követően rátért a Scholia nevű projekt bemutatására, mely a Wikidata adatait használja fel, hogy bibliográfiai információkat szerezzünk a tudományos élet szereplőiről, szerzőkről vagy intézményekről.

Az előadás haszna, hogy egy olyan kezdeményezést ismerhetünk meg, mely még korántsem befejezett − hiszen még a legnagyobb, s legbefolyásosabb kiadók nem csatlakoztak hozzá, hogy példaként csak az Elseviert említsük −, viszont előremutató lépést tesz a mindenki által elérhető tudás felé.

ꕥ ꕥ ꕥ ꕥ ꕥ

James Hendler: The Semantic Web: vision, reality and revision

A Szemantikus Háló: elképzelés, valóság és újragondolás
(James Hendler mesterséges-intelligencia kutató a Rensselaer Polytechnic Institute-nél [USA], valamint a Szemantikus Háló egyik megalkotója.)
https://www.youtube.com/watch?v=FchE3ktj7U0&feature=youtu.be

Bevezetőjében Hendler elmondja, hogyan ismerkedett meg Tim Berners-Lee által a szemantikus háló, valamint a tudásdiagram fogalmakkal. Elmondja, hogy alakult ki a szemantikus háló projekt: kezdődött a DARPA jelölőnyelvvel[11], ebből nőtt ki aztán az RDF[12], majd az OWL[13] nyelv. Megjelentek szemantikus keresők; a kezdetben óriási ver-

senyből mára egy igazi motor maradt: a Google. A cég 2009-ben megjelent cikkében[14] arról írt, hogy csak az adatra van szükségünk, a szemantika nem számít. Ehhez képest, meglepő módon, 2016-ban már arról számol be ugyanez a cég, hogy keresései 40%-ban mutatnak találatot beágyazott metaadatokra. Azaz bebizonyosodott, hogy az ipar igenis nagymértékben használja a szemantikus hálót.

Jelenleg kutatások sora folyik annak érdekében, hogyan javítsuk a hibákat. Erre hoz Hendler két szemléletes példát 2014-ből a Yahoo nevű cég keresései közül. Az egyik, amiben az Ice Cube nevű rapperre kerestek: adatlapján ott a fényképe, a metaadatok, amik meglehetősen pontosak, azonban a kép alatti leírás kis, kocka alakú, fagyott halmazállapotú vízről szól. Másik példája a Michelangelo-ra indított keresés: ez esetben a pontos életrajzi adatok mellé egy tini nindzsa teknőcöt ábrázoló kép került. Nincsen ugyanis metaadat, nincs semmi, ami azt mondaná a számítógépnek, hogy a kép egy teknőc képe.

Hendler felveti az együttműködési problémát: a fontos az lenne, hogy az emberek tudjanak kapcsolódni egymás dolgaihoz, és hogy a kapcsolatok minél leíróbbak legyenek. A linkeknek segíteniük kellene megérteni, hogy miért történt és mit jelent, ha két dolgot összekapcsoltak.

Ez után kitér a DIVE-ra: „fejesugrás" az adatokba. Ez egy betűszó: *Discover* (felfedezés), *Integrate* (egységesítés), *Validate* (megerősítés), *Explain* (megmagyarázás). Ezek voltak az eredeti eszmék a szemantikus háló mögött. Magának a hálónak a célja pedig az lenne, hogy mások tudjanak arról, milyen adataink vannak, és azt is, hogyan férhetnek hozzá. A metaadatok összeállítása helyett az összekapcsolt adatok, tudásdiagramok építése került előtérbe. Azonban itt lenne az ideje, hogy visszatérjünk az eredeti elképzeléshez, mert az igazán érdekes a metaadatok lehetséges felhasználása az összekapcsolt adatokban.

Előadásának végén pedig szól néhány szót egy új, az MIT Press és a Kínai Tudományos Akadémia közös gondozásában induló online folyóiratról[15], mely felhasználható adatokkal foglalkozik többféle területen – így a metaadatokkal is könyvtárak és különböző gyűjtemények részére.

ꕥ ꕥ ꕥ ꕥ ꕥ

George Oates: Every Collection is a Snowflake

Minden gyűjtemény az egyedisége és összekapcsolódási lehetőségei alapján egy hópehelyhez hasonlítható
(George Oates szoftverfejlesztő 2008 óta dolgozik a kulturális örökség ágazatában; 2014-ben alapította meg a londoni székhelyű Good, Form & Spectacle cégét, amelynek célja explicit módon bemutatni, hogyan lehet a medaatadatok feltárásához felhasználni az emberközpontú tervezést és a gyors szoftverfejlesztést)
https://youtu.be/5JAPbrvkwmc

George Oates az előadás bevezetőjében elmondta, hogy kezdetben nem volt egységes a gyűjtemények katalógusainak összeállítása, valamint az adatokat nem kapcsolták össze egymással, így azok gazdagsága sem tudott megfelelően tükröződni. Az utóbbi évtizedekben a webre feltöltött digitális anyagainkat is a világ számos nyelvén, változatos részletességgel írjuk le. Az előadó a közgyűjteményi intézetekkel való együttműködések során arra törekedett, hogy az adatok inkonzisztenciájának megmutatása helyett új szemszögből való rálátást mutasson a gyűjteményekre az adatok létrehozói számára.

A **Flickr** képmegosztó szolgáltatásnál és a közintézmények számára fejlesztett **Flickr Commons** programnál kihangsúlyozta a felhasználók aktív részvételét a metaadatolásban: a képekhez szabad szövegezésű tag-eket hozzáadva segíthették azok kategóriákba való besorolását.

Az **Open Library Project**, egy Wikipedia alapú, szerkeszthető e-könyvkatalógus vezetőjeként kiemelte, hogy létrehoztak egy olyan interface-t, ahol a felhasználóknak lehetősége volt szerzői nevek egybevonására. A legnagyobb feladatnak az bizonyult, hogy összegyűjtsék az adott kiadású könyvekhez tartozó azonosítókat (pl. ISBN, Internet Archive, WorldCat azonosítók).

A **Two Way Street** elnevezésű fejlesztést a British Museum kb. 2,2 millió katalógusrekordján végezte 3 ember 1 hét alatt. Megfelelő vizualizációs technikákkal meg tudták mutatni többek között az egyes tételek beszerzési információit, anyagát, lelőhelyét stb.

Az előadó végül az OCLC kutatókönyvtárak csoportjának tanulmányát alapul véve azt javasolta a közintézményeknek, hogy elemezzék a leggyakrabban használt és az esetlegesen felesleges

MARC-mezőket. George Oates további tanácsa, hogy érdemes alaposan körüljárni az adott intézmény katalógusának felhasználói közönségét és felmérni, hogy vajon megkapják-e mindazt, amire szükségük van.

A bemutatott tapasztalatok jó példák voltak arra, hogy egy gyűjteménynél hogyan könnyíthető meg a keresési folyamat, ha az adatokhoz interface-t, vizualizációs technikát és hiperlinkeket adunk. Minden gyűjtemény egyedi, megismételhetetlen, mint ahogy nincs két egyforma hópehely sem. Ha össze tudnak kapcsolódni, egymásra épülni, a keresési szempontok soha nem tapasztalt mértékű gazdagságát figyelhetjük meg.[16]

Az összefoglalók szerzői

Perlaki Gabriella <gabriella.perlaki@ek.szte.hu>
Kiss Márta Éva <marta.kiss@ek.szte.hu>
Szügyi-Szűcs Judit <judit.szucs@ek.szte.hu>
Kiss Zsuzsanna <zsuzsanna.kiss@ek.szte.hu>
Kovács Anita <anita.kovacs@ek.szte.hu>
Laskay Krisztina <krisztina.laskay@ek.szte.hu>
dr. Kokas Károly <karoly.kokas@ek.szte.hu>
– az SZTE Klebelsberg Könyvtár (Szeged) munkatársai

Absztraktok

Szemantikus web és könyvtár: egy konferenciasorozat margójára *(Hat előadás a SWIB 2017/2018-as konferenciáiról)*

Jó néhány éve már, hogy világossá vált, hogy a jövő egyik legjelentősebb könyvtárosokra váró feladata információs rendszereink webtérbe és webtérrel való integrációja lehet. A könyvtárak és állományaik, leíró adataik így az óriási hipertext-tér részei lesznek. A könyvtári katalogizálás újfajta szemlélete lehetővé teszi, hogy az egyes könyvtárakban leírt objektumok összekapcsolhatók legyenek, mégpedig szemantikus módon az egész hálózati hipertext-térrel és ilyen módon a könyvtári adatfeltárás maga is részévé válik ennek a webtérnek, mintegy a webtérbe beágyazva (a „linked data” technológia könyvtári vetülete). Ma már a könyvtári területen is megjelentek a konkrét szemantikus webes fejlesztések. Mindezekről az egyik legfontosabb hely a mai Európában, ahol „élőben” tájékozódhatunk, az ún. SWIB konferencia-sorozat. A dolgozat a legutóbbi két (2017 és 2118) konferenciáról emel ki hat előadást, hogy konkrétan is bemutassa őket, a konferenciasorozat videó-archívumának felhasználásával.

Hivatkozások

1 Az egyik legutóbbi és nagyon alapos cikk a témában, amit Fülöp Endre 2018-as sikeres networkshop-os előadása nyomán írt. Fülöp Endre: A szemantikus háló két fogalma, a katalógusok új generációja és a könyvtárak szerepe= TMT, Vol 65, 7-8 (2018) – http://tmt.omikk.bme.hu/tmt/article/view/7154

2 http://swib.org/swib17/ és http://swib.org/swib18

3 https://public.oed.com/appeals/words-where-you-are/ A felhívás szövege és a hashtag #wordswhereyouare [2019.02.18.]

4 https://www.libcrowds.com/collection/playbills [2019.02.18.] Int he Spotlight projekt honlapja.

5 http://data.nationallibrary.fi/bib/me/CFENNI – a FENNICA honlapja [2019. 03. 01.]

6 http://data.nationallibrary.fi/ - a Finn Nemzeti Könyvtár Open Data szolgáltatása [2019. 03. 01.]

7 http://annif.org/ – Annif honlapja [2018.03.11.]

8 https://finna.fi/?lng=en-gb – Finna honlapja [2019.03.11.]

9 http://m.annif.org/ – mobil web applikáció elérése [2019.03.07.]

10 https://github.com/NatLibFi/Annif – Annif dokumentációja és a tesztelések eredményei [2019.03.07.]

11 https://en.wikipedia.org/wiki/DARPA_Agent_Markup_Language

12 http://www.w3c.hu/forditasok/RDF/REC-rdf-concepts-20040210.html

13 http://www.w3c.hu/forditasok/OWL/REC-owl-features-20040210.html

14 https://static.googleusercontent.com/media/research.google.com/hu//pubs/archive/35179.pdf

15 www.data-intelligence-journal.org

16 Hubay Miklós Péter: A BIBFRAME és a könyvtári feldolgozás új keretei. [Szakdolgozat]. Eszterházy Károly Főiskola, Eger (2016) – http://mek.oszk.hu/15600/15678/

(Kokas Károly
SZTE Klebelsberg Könyvtár)