

A SZEMANTIKUS WEB ÚJ LEHETŐSÉGEK ÉS KIHÍVÁSOK KÖNYVTÁRI KÖRNYEZETBEN

Bevezetés

„A szemantikus web nem egy önálló web, hanem a jelenleginek a kiterjesztése, ahol az információknak jól definiált jelentést adunk, amely elősegíti a számítógépek és az emberek közti együttműködést. Az első lépéseket már megtettük abban az irányban, hogy a jelenlegi web struktúrájába beleszőjük a szemantikus webet. A közeli jövőben ezek a fejlesztések jelentős új funkciókat kaphatnak, ahogy a gépek egyre alkalmasabbak lesznek azon adatok feldolgozására és „megértésére”, amelyeknek ma csupán a megjelenítésére képesek.”¹ A fenti sorok íróját, Tim Berners-Lee-t bátran sorolhatjuk a 20. század nagy álmodói közé. Az információhoz való egyetemes és gyors hozzáférés lett a 21. század emberének „Bibliotheca Universalis”-a, amelyhez fontos lépést jelenthetnek a fenti sorok, és amelyhez egy – Paul Otletől a Memexen és a Xanadun keresztül a World Wide Webig vezető – nagy ívű történelmi fejlődés jelenti a hátteret.

Az univerzalizmus ezen heroikus kísérletei a nagy közös katalogizálási projektek, illetve az azokra épülő könyvtárközi dokumentumellátó rendszerek nyomán mára összeálltak egy globális hálózati infrastruktúrává, amely a legkülönbözőbb hordozókon tárolt információkhoz kínálnak egyetemes hozzáférést. Tanulmányomban a webes szolgáltatások újabb generációja által vázolt lehetőségeket és kihívásokat tárgyalom könyvtári szempontból a magyarul és angolul megjelent szakirodalom alapján a digitális könyvtár kontextusában.

A digitális könyvtár

Mi különbözteti meg a digitális könyvtárat a kibertér egyéb pontjaitól? A digitális könyvtárak kapcsán a tévképzetek két fő csapásirány mentén csoportosulnak. Egyrészt felmerül, hogy a web és a hipertext olyannyira megváltoztatta az információ fogyasztásának mikéntjét, hogy a könyvtárak sem állhatnak meg régi önmagukban; másrészt pedig hogy a web egyenesen könyvtári funkciókat képes ellátni. A tévképzetek eloszlata és néhány markáns definíció áttekintése során kivilágosodik a digitális könyvtárak természete.

„A digitális dokumentumok mindig hibridek és illékonyak: a globális információs infrastruktúra olyan metaarchívumot feltételez, amelyben az őrzésre kerülő bitek folyamatosan átrendeződnek, elannyira, hogy a permanens átalakulás szinte maga a

¹ Berners-Lee, T. 2001.

lényeg. (...) A mindent háttérbe szorító folyamatos jelenidő nem csekély megrázódatást okoz mindazon intézményeknek, amelyek mindig is a történelmenkívülséget, a társadalmi idő felett állást képviselték.”² – írja György Péter. A szövegek lezáratlansága és a folytonos jelenidő azonban nem érvényes a könyvtári dokumentumokra, ugyanakkor a World Wide Weben elérhető dokumentumok tömege sem tekinthető homogén közegnek még akkor sem, ha a hivatkozások, csatolások, ugrópontok bonyolult rendszere, a hiperszöveg belső törvényszerűségei ezt sugallják. Hangsúlyoznunk kell tehát a digitális könyvtár **lehatároltságát** az interneten elérhető dokumentumok sűrűjében.

És ezt a határozottan elkülöníthető részt mi teszi könyvtárrá a környezetéhez képest? Azaz a World Wide Web önmagában miért nem tekinthető könyvtárnak? Deegan és Tanner szerint „A web nem egy tudatosan menedzselte környezet, nincsenek gyűjteményszervezési elvei, és mindenek között a legfontosabb, a digitális objektumokat nem tekintik tartós értéknek, annak ellenére, hogy némelyek azok.”³ A könyvtárban tárolt dokumentumok **tartóssága**, a gyűjtemény **tudatos menedzselése** alapvető ismérvek mind hagyományos, mind pedig digitális könyvtári környezetben.

Lagoze és Payette szerint a digitális könyvtár „digitális objektumok és szolgáltatások gyűjteménye, amely szolgáltatások támogatják az objektumok tárolását, feltárását, (vissza)keresését és megőrzését.”⁴ – a **szolgáltatások** megléte szintén meghatározó ismérvek.

A digitális könyvtárak négy fő karakterjegye Deegan és Tanner nyomán:

- A digitális könyvtár digitális objektumok szervezett gyűjteménye;
- A digitális objektumokat gyűjteményszervezési elvek mentén készítik, vagy gyűjtik össze;
- A digitális objektumokat összefüggésükben teszi elérhetővé, a megfelelő szolgáltatásokkal, amelyek lehetővé teszik az erőforrások visszakeresését és használatát, éppen úgy, mint tennék azt bármelyik könyvtári dokumentummal;
- A digitális objektumokat hosszú távon is elérhető erőforrásokként kezelik és a megfelelő eljárások mentén biztosítják a minőségüket és a fennmaradásukat.⁵

Lagoze és kollégái az egyetemi könyvtárak funkcióváltására utalva – a digitális könyvtárak szélesebb kontextusban való értelmezését hangsúlyozzák. Reeves⁶ azon megállapításából indultak ki, hogy a multimédia és egyéb információtechnológiai eszközök igazi ereje abban rejlik, hogy egy interaktív tanulási környezetet képesek teremteni a papír alapú dokumentumok alapján való tanulás egyszerű befogadói attitűdje helyett. A digitális könyvtáraknak az információt kontextusában kell megjeleníteniük, folyamatosan követve a használók „lélegzetvételt” is, akik a különböző

² György P. 2002:149.

³ Deegan, M.–Tanner, S. 2002:26.

⁴ Lagoze, C.–Payette, S. 1998.

⁵ Deegan–Tanner, 2002. 22. p.

⁶ Reeves, T. C. 1998.

felületeken megjelenő tanulási környezet aktív alakítóivá is válnak egyben.⁷ Ez a megközelítés a hagyományos könyvtárak **szociális és tanulási interakciós tér** funkcióját is átmenti digitális környezetbe. Ez utóbbi elképzelések legösszefogottabban formában eddig a *könyvtár 2.0* koncepcióban kristályosodtak ki.⁸

A rövid kitekintés végén eljuthattunk Crawford és Gorman azon megállapítására, miszerint a könyvtárak funkciója, történelmi szerepe lényegében ugyanaz maradt a legutóbbi évtizedek rohamos technikai fejlődését követően is: begyűjteni, tárolni és visszakereshetővé tenni a tudást és az információt.⁹ A digitális könyvtárat a tárolt dokumentumok lehatároltsága, tartóssága, a gyűjtemény tudatos menedzselése, szolgáltatások fenntartása és a virtuális találkozóhely jellege különbözteti meg az információs tér egyéb pontjaitól.

A felhasználó

Milyen igények fogalmazódnak meg a keresleti oldalon? A digitális kor felhasználóját az interakció, a szórakozás, a társas tevékenységek és a versenyzés, a versenyszerűség különbözteti meg a régmúlt könyvtárhasználójától,¹⁰ amely a könyvtáros szakma számára egy aktívabb, ugyanakkor periférikusabb szerepet jelöl ki, aki rendezi, felépíti, útjelzőkkel látja el az információs teret, abban kapcsolatokat definiál, és a szociális funkciónak eleget téve segíti a használót az eligazodásban.¹¹

A szemantikus web

A szemantikus web a world wide web kaotikusságára adott válaszként, annak technológiai úton való rendezésére tett kísérlet, metaadatok és a formális logika segítségével felépített keresési lehetőségekkel. Célja egy olyan metaadat-infrastruktúra létrehozása, amely lehetővé teszi a Weben lévő adatok integrálását, a közöttük lévő kapcsolatok definiálását és jellemzését, illetve az adatok értelmezését.

Szerkezete a következő rétegekkel írható le:

- metaadatok: az erőforrások és tulajdonságok leírása;
- sémák: a fogalmak hierarchikus leírása;
- logikák: leíró logikák.¹²

Az egyes rétegekhez különböző nyelveket, specifikációkat hoztak létre, amelyek egy része specifikus a szemantikus webre más részük nem. Az adatok egyértelmű címezhetőségét a weben az URI-k (Universal Resource Identifier) biztosítják. A dokumentumok így önleírók lesznek. Az önleíró dokumentumnak egységes szintaktikájúnak is kell lenni. Ezt biztosítja az XML (eXtensible Markup Language) nyelv.

⁷ Lagoze, C. et al. 2005.

⁸ Miller, P. 2006., Maness, J. M. 2006.

⁹ Crawford, W.–Gorman, M. 1995:7.

¹⁰ Nicholas, D. et al. 2000.

¹¹ Tóth M. 2003.

¹² Lu, S. et al. 2002.

Szükség van egy precíz adatmodellre, amely formális keretet ad az adatok egymáshoz való kapcsolatainak definiálására, és a kapcsolatok leírására. Erre szolgál az RDF nyelv¹³ (Resource Description Framework) Az adatok közti kapcsolatokat, illetve a kapcsolt adatok referenciáit, el kell tudni érni, le kell tudni kérdezni (SPARQL)¹⁴ A kapcsolatok leírására szolgáló terminológiát definiálni kell tudni (RDFS, OWL, SKOS¹⁵). A kapcsolatokon, illetve azok leírásán, logikai következtetéseket kell tudni levonni (OWL, RIF).¹⁶ Ezek egy része szabványként létezik, a más részük a fejlesztés stádiumában van. Ezen specifikációk határozzák meg a szemantikus web lépcsős modelljét¹⁷, amely az önleíró dokumentumra épülve a metaadatok segítségével, a formális logikai szabályok szerinti következtetések levonására lesz képes.

A szemantikus web könyvtári trendjei

Milyen hatással lehet ez az új webes infrastruktúra a digitális könyvtárak fejlődésére, a felhasználók attitűdjeire, illetve a szolgáltatási modellre? A válasz megadásához kizárólag a jelen technológiai környezetéből indulhatunk ki. A szemantikus web az információ legegységesebb hordozó egységeitől kiindulva a rendszer teljes komplexitásáig változásokat hordoz magában. A következőkben három lehetséges trendet emelek ki: egy a feldolgozás, egy az információ-visszakeresés, egy pedig a tájékoztatás területén jelentkeznek.

Szemantikus web környezetben az információ **feldolgozása** URI-kről tett állításokat jelent. Mivel egy URI, azaz a szemantikus web legkisebb alkotóeleme nem egy dokumentumhoz, hanem egy adathoz, egy információhoz kapcsolódik, a dokumentumnál mélyebb a feldolgozottság szintje, így a szemantikus web képes lesz a „dokumentumok belsejéből” is választ szolgáltatni kérdéseinkre.

A lekérdezés, az **információ-visszakeresés** az URI-kről tett állítások alapján lehetséges. Az RDF nyelv hármas szerkezete olyan állítások lekérdezését is lehetővé teszi, amelyet soha nem tettek, csak logikusan következnek a többiből. A könyvtári és nem könyvtári forrásból származó URI-k, illetve állítások integrálhatók, így a web egyre inkább egyetlen nagy szövegre fog emlékeztetni, amelynek részét képezik a digitális könyvtár által közzétett és kezelt dokumentumok. A felhasználó oldaláról így folytatódik az a tendencia miszerint a digitális könyvtár egyre kevésbé különböztethető meg a kibertér egyéb pontjaitól.¹⁸ A könyvtár aktivitása a periférián ugyanakkor tovább nő, hiszen láthatatlanságában is őt terheli a megbízható információ szolgáltatásának felelőssége.

¹³ Az RDF/XML szintaxis specifikációja

<http://www.w3c.hu/forditasok/RDF/REC-rdf-syntax-grammar-20040210.html>

¹⁴ SPARQL Query Language for RDF

<http://www.w3.org/TR/2006/CR-rdf-sparql-query-20060406/>

¹⁵ SKOS Core Guide <http://www.w3.org/TR/2005/WD-swbp-skos-core-guide-20051102/>

¹⁶ Az OWL Web Ontológia Nyelv – Áttekintés

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

¹⁷ Fülöp Cs. et al. 2004.

¹⁸ Moyo, L. 2004.

Tovább nő az internet szerepe a **tájékoztatásban**. A szemantikus web a jelenlegi keresőrendszereknél hatékonyabb eszközt kínál arra, hogy referenzkérdésekre a világhálón tárolt információk alapján adjunk választ. Mindezek mellett erősödik az először egy évtizeddel ezelőtt leírt folyamat, miszerint a könyvtári referenzkérdések egyre bonyolultabbakká válnak, mert az egyszerűbb kérdésekre a hálón könyvtárosi segítség nélkül is találnak választ a használók.¹⁹

A szemantikus web technológiák könyvtári kiaknázása óriási lehetőséget jelent a jövőre nézve, erre vonatkozóan számos cikk jelent meg a hazai szakirodalomban az egyes technológiák kapcsán.²⁰ A szemantikus keresésre épülő tájékoztatási eszközök megjelenése ezekhez képest egy újabb stádiumot jelenthet.

Míg ezen fenti megállapítások csak egy részben elkészült technológia várható hatásának függvényében nyerhetnek bizonyosságot, a szakma előtt álló rövid távú kihívások világosan kirajzolódtak már a szemantikus web fejlesztések kezdeti stádiumában.

Rövid és hosszú távú kihívások

Mint más technológiák esetében is, a fejlesztés igazi mozgatórugója maga a lehetőség, amelyet az új technológia magában hordoz.²¹ Mi lehet a könyvtáros szakma szerepe az új generációs webalkalmazások megvalósításában? Tekintve, hogy a technológia megvalósítása egy eddig soha nem látott méretű metaadat infrastruktúrát feltételez, adódik a könyvtáros szakma ezirányú felelőssége. Miként fordíthatja szakmánk mindezt a szolgáltatások színvonalának javítására?

Rövid távon szemantikai átjárhatóságot kell biztosítani a szélesebb web környezettel: a szemantikus web technológiákkal kompatibilis módon kell végezni a források közzétételét és a digitális (illetve digitalizált) dokumentumok feldolgozását. Hatékony osztályozási és indexelési mechanizmusokat kell kidolgozni a digitális könyvtárak egyes típusaira.²²

Az első pillanattól kezdve reagálni kell a környezet változására, amely hosszabb távon az új technológiai lehetőségeket, és a felhasználók arra adott választát, megváltozott attitűdjeit tükröző szolgáltatási paradigma felrajzolását teszi szükségessé.

Összegzés

A szemantikus web technológiák fejlesztése a digitális könyvtár koncepciók fényében előrevetíti a könyvtárautomatizálás időszakában elindult tendenciák erősödését az információ legegyszerűbb hordozó egységeitől kiindulva a rendszer teljes komplexitásáig: a feldolgozástól a tájékoztató munkáig. Rövid távú kihívásokat jelent a szemantikus átjárhatóság biztosítása a dokumentumok feldolgozása során,

¹⁹ Bushallow-Wilber, L. et al. 1996.

²⁰ Ld. Bíró Sz. 2004., Fülöp Cs. et al. 2004., Salgáné M. M. 2004., Ungváry R. 2004.

²¹ Lu S. et al. 2002.

²² Goossens, P. 2003., Menon, B. 2003.

illetve hatékony indexelési és osztályozási mechanizmusok kidolgozása. Hosszú távon pedig egy új kihívásokra reflektáló szolgáltatási paradigma felrajzolása a cél.

Irodalomjegyzék

- Berners-Lee, T. et al. 2001. *The Semantic Web*. Scientific American, May 17, 2001.
http://www.ciam.com/print_version.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21
- Bíró Sz. 2004. *A szövegfeldolgozás modern eszközei – az SGML és XML nyelvek*. Tudományos és Műszaki Tájékoztatás, 51./2.
http://tmt.omikk.bme.hu/show_news.html?id=3541&issue_id=448
- Borgman, C. L. 2000. *From Gutenberg to the global information infrastructure: access to information in the networked world*. Cambridge, MA, London: MIT Press.
- Bushallow-Wilber, L. et al. 1996. *Electronic Mail Reference Service: A Study*. RQ 35./3.
- Crawford, W. – Gorman, M. 1995. *Future Libraries, Dreams, Madness and Reality*. Chicago, London: American Library Association.
- Deegan, M. – Tanner, S. 2002. *Digital futures, Strategies for the information age*. London: Library Association Publishing.
- Fülöp Cs. et al. 2004. *A metaadatsémák és a szemantikus web: egységesítés és specializáció a metaadatok világában*. Tudományos és Műszaki Tájékoztatás, 51./7.
http://tmt.omikk.bme.hu/show_news.html?id=3648&issue_id=452
- Goossens, P. 2003. *ELAG 2002: report of a library systems seminar on the Semantic Web*. Program: electronic library and information systems, 37./4.
- György P. 2002. *Memex A könyvbe zárt tudás sorsa a 21. században*. Budapest: Magvető kiadó.
- Lagoze, C. – Payette, S. 1998. *An infrastructure for open-architecture digital libraries*.
<http://ncstrl.cs.cornell.edu/DienstUI/1.0/Display/ncstrl.cornell/TR98-1690>
- Lagoze, C. et al. 2005. *What Is a Digital Library Anymore, Anyway? Beyond Search and Access in the NSDL*. D-Lib Magazine, 11./11.
<http://www.dlib.org/dlib/november05/lagoze/11lagoze.html>
- Lu, S. et al. 2002. *The Semantic Web: opportunities and challenges for next-generation Web applications*. Information Research, 7./4.
<http://InformationR.net/ir/7-4/paper134.html>
- Nicholas, D. et al. 2000. *The information 'player': a new concept for the information user*. Aslib Proceedings, 52./5.
- Maness, J. M. 2006. *Library 2.0 Theory: Web 2.0 and Its Implications for Libraries*. Webology, 3./2. <http://www.webology.ir/2006/v3n2/a25.html>
- Menon, B. 2003. *Le web semantique: de nouveaux enjeux documentaires?* Documentaliste, 40./6.
- Miller, P. 2006. *Coming together around Library 2.0 A Focus for Discussion and a Call to Arms*. D-Lib Magazine, 12./4. <http://www.dlib.org/dlib/april06/miller/04miller.html>
- Moyo, L. 2004. *Electronic libraries and the emergence of new service paradigms*. The Electronic Library, 22./3.
- Reeves, T. C. 1998. *The Impact of Media and Technology in Schools: A research Report prepared for the Bertelsmann Foundation*.
<http://it.coe.uga.edu/~treeves/edit6900/BertelsmannReeves98.pdf>
- Salgáné M. M. 2004. *Az XML: új perspektívák a könyvtár-informatikában*. Tudományos és Műszaki Tájékoztatás, 51./2.
http://tmt.omikk.bme.hu/show_news.html?id=3541&issue_id=448

- Szeredi P. et al. 2005. *A szemantikus világháló elmélete és gyakorlata*. Budapest: Typotex kiadó.
- Tóth M. 2003. *A felhasználó mint játékos, a könyvtáros szakma szerepváltása a digitális korban a felhasználói igények mentén*. Tudományos és Műszaki Tájékoztatás, 50./2. http://tmt.omikk.bme.hu/show_news.html?id=1668&issue_id=48
- Ungváry R. 2004. *Tezaurusz és ontológia avagy a fogalmi ismertetőjegyek generikus öröklődésének formalizálása*. Tudományos és Műszaki Tájékoztatás, 51./5. http://tmt.omikk.bme.hu/show_news.html?id=3615&issue_id=450