



A könyvtár a számítógépes hálózatok világában

Egyre inkább érzékelhető, hogy az amerikaiaknak csökken a társadalmi és polgári elkötelezettsége, mind kevésbé vesznek részt közösségi akciókban, vállalkozásokban, szavaznak, lépnek be pártokba. Ezt a jelenséget lehet az ingázásnak, a fokozódó munkahelyi elvárásoknak, a technológia fejlődésének is betudni, mindenesetre megélnék a viták az ún. társadalmi tőke fogalmáról és használatáról. A könyvtárosok arról is panaszkodnak, hogy a vitázók nem tekintették a könyvtárat a társadalmi tőke példás építőjének.

Ez részben érthető is, hiszen az utóbbi harminc évben a könyvtárakat főként a számítógépesítés kötötte le. A számítógép a szolgáltatás standard elemévé vált: 1996-ban a szolgálati pontok 44,6%-a kapcsolódott az internethez, ma pedig 100%-uk. 14 millió lakos használja rendszeresen a közkönyvtárak számítógépeit, hogy hozzáférjen az internet szolgáltatásaihoz.

Az internet nagykorúvá vált. Világszerte mind több és több ember tölt legalább valamennyi időt az interneten kapcsolatot keresve másokkal, kutatva, vásárolva, alkotva, publikálva. Más szóval úgy is mondhatnánk, hogy az internet-hozzáférés abban segíti az embereket, hogy megkeressék és megtalálják önmagukat. A webalapú technológiák teljes mértékben áthatották a kereskedelmi, kutatási, sőt a személyes interakciókat is. Az elektronikus levelezés, a listák, a blogok, a wikik és a társadalmi hálózatok, azaz a társadalmivá vált háló arra biztat mindenkit – nem csak a könyvtárakat! – hogy próbálja meghatározni és kifejezni helyét az új média világában.

Az internet tehát beépült az emberek mindennapjába. Az OCLC felmérése szerint azonban – bár minden internetes tevékenység növekedést mutatott fel 2005 és 2007 között –, csökkent a könyvtári webhelyek használata.

Ezzel szemben a könyvtárak fizikális használata szerény mértékben növekedett. Nem vált be az a

jóslat, hogy az internet növekvő használata a könyvtárhasználat csökkenéséhez vezet. Az ALA (American Library Association) jelentése szerint 1994 és 2004 között a közkönyvtári látogatások száma 61%-kal, a kölcsönzések száma 28%-kal növekedett. A könyvtáraknak továbbra is egyedülálló funkciója az, hogy az ismereteket és az információkat keresőket felkészült és képzett személyzettel segítsék.

Az internethasználó amerikai felnőttek aránya az 1995-ös 15%-ról mára 73%-ra nőtt. A 18-19 éves korosztályban 90%-ra tehető az online hozzáférést igénylők csoportja.

A növekvő fizikális használat azonban nem nyugtathatja meg a könyvtárosokat, mert alighanem nem tettek eleget az online kapcsolatok kiépítése és a külső kapcsolatok ápolása terén. Elhanyagolták tudomásul venni az információhoz való hozzáférés megszervezésében betöltött szerepüket úgy, mint elsődleges társadalmi szerepet. Ha a könyvtáros nincs jelen mint a közösség aktív építője saját, közvetlen környezetében, akkor bizony a könyvtáron kívüli területekről és az információ terekről is hiányzik, amelyek pedig mélyen beágyazódtak ma már a mindennapi életbe.

A dokumentumoknak ugyanis „társadalmi életük” is van. Nemcsak az információs tér strukturálásában és megszervezésében segítenek bennünket, hanem a társadalmi tér megteremtésében és kezelésében is. A dokumentumok körül csoportok keletkeznek és viták folynak, s így a dokumentum fogalmát ki kell bővíteni a körülötte folyó társadalmi interakciókkal is. Ha megfeledekezünk a dokumentumok – és használóik! – társadalmi vonatkozásairól, akkor megfeledekezünk saját szerepünk társadalmi jellegéről is.

A trend megfordítható és visszatérhetünk alapértékeinkhez, hozzáigazítva őket a jelen információs színteréhez és gyakorlatához. A kérdés nem annyira az, hogy hol a könyvtár helye az információs

térképen, mint inkább az, mi is a könyvtár a mai világban.

A könyvtár összetevő elemei az állomány, a hely, a szolgáltatás, a műveletek és az értékek. A digitalizálódás előrehaladtával az állományok és a szolgáltatások a falakon kívülről is hozzáférhetővé válnak. A hely, amit egykor a könyvtár jelentett, egyre inkább a felülvizsgálat tárgyává válik. Új fontosságot nyernek viszont a szolgáltatások, az értékek és a személyközi műveletek.

A számítógépesítéssel való megszállottságunk ugyan megmagyarázza, miért nem kerültünk a

társadalmi elkötelezettségről folyó viták középpontjába, de paradox módon visszavezethet bennünket a közösségépítés központi szerepébe. Visszavisz minket a könyvtárosság alapelveihez, és egyértelműen hozzáköt a hagyományos könyvtári szolgáltatásokhoz, amelyek a múlt századforduló óta a helyi közösségeket építették. Ezt kell tennünk ma is, könyvtáron belül, könyvtáron kívül és online egyaránt.

/HILL, Chrystie: *Inside, outside and online.* = *American Libraries*, 40. köt. 3. sz. 2009. p. 38–41./

(Papp István)

Egy digitális könyvtár használhatóságának és hasznosságának értékelése

Bevezetés

A skót *University of Strathclyde* kutatói egy klinikai döntéseket segítő e-könyvtári szolgáltatást tesztelttek orvosokkal és más egészségügyi szakemberekkel, hogy megállapítsák, mennyire felel meg a felhasználói elvárásoknak.

Az információs rendszerek *használhatósága* és *hasznossága* két kulcsfogalom, melyek közösen határozzák meg azt, hogy az emberek mennyire elégedettek egy rendszerrel és mennyire használják. Előbbi az ember-gép kapcsolatra, elsősorban a kezelőfelület jellemzőire vonatkozik, utóbbi viszont arra, hogy a rendszer kielégíti-e a használója információs igényeit. A szakirodalomban többnyire a digitális könyvtári szolgáltatások – már elég jól kialakult elvek mentén történő – használhatósági vizsgálatairól lehet olvasni, de egyre többen javasolják ennek a két jellemzőnek az együttes kezelését és mérését, mert hiába jól használható egy forrás, ha nem nyújt releváns információt a keresőknek. Az igazi kérdés persze az, hogy pontosan mit mérjünk és hogyan, amikor meg akarjuk állapítani, hogy egy információforrás mennyire hasznos a felhasználóknak?

Használhatóság

A használhatóság értelmezésére többféle javaslat is született már. A szoftverek *ISO 9126-1* szabványa a következő jellemzőket sorolja fel ezzel kapcsolatban: érthetőség, megtanulhatóság, kezelhe-

tőség és vonzerő. Az ergonómiai követelményekre vonatkozó *ISO 9241-11* specifikáció viszont az eredményességet, a hatékonyságot és a felhasználói elégedettséget említi. A szakirodalomban ezeken kívül előfordulnak még egyéb fogalmak is: memorizálhatóság, hibaarány, biztonságérzet, egyszerűség, esztétikai hatás, navigálhatóság és terminológia. A jelen kutatás során az alábbi paraméterek segítségével mérték a rendszer használhatóságát:

- *eredményesség* (mennyire sikerül a keresési feladatokat megoldania a felhasználónak?);
- *hatékonyság* (milyen gyorsan és könnyen sikerül egy-egy feladatot megoldani?);
- *esztétikai hatás* (kellően következetes és a célnak megfelelő-e a felület kialakítása?);
- *navigáció* (milyen könnyen tud a felhasználó mozogni a rendszerben és nem érzi-e magát „elveszve”?);
- *terminológia* (érthetők-e mindenkinek a rendszerben előforduló szakkifejezések, és mennyire következetes, illetve logikus ezek használata?);
- *tanulhatóság* (milyen gyorsan tudja elkezdni egy új felhasználó a rendszer használatát, illetve újratanulni azt később, amikor visszatér?).

Hasznosság

Az információforrás által kínált tartalom és szolgáltatások, valamint a felhasználói igények és elvárások illeszkedésének mértéke határozza meg a hasznosságot. Ennek a megállapításához is különféle – bár a használhatóságnál kevésbé változatos

– jellemzőket javasoltak az ezzel foglalkozó szakemberek. Például: érték, megbízhatóság, érvényesség, pontosság, relevancia, formátum, szint és lefedettség. Ezek közül a következőket választották ennél a kutatásnál:

- *relevancia* (mennyire alkalmas a megtalált információ a felhasználó igényének kielégítésére, illetve megfelelő válaszokat ad-e a rendszer a keresőkérdésekre?);
- *megbízhatóság* (kellően pontos, hiteles és elmentmondásmentes-e az információ?);
- *érvényesség* (mennyire friss és aktuális a kinyert információ, nem avult-e már el?).

Természetesen vannak átfedések és függőségek ezek között az attribútumok között, de mindegyiknek van elég egyedi jellegzetessége és önállóan is mérhető.

Módszertan

A használhatóság értékelése történhet fejlesztő (*formative*) vagy összegző (*summative*) szinten, és rendszerint valamilyen vizsgálattal (felhasználók részvétele nélkül) és/vagy teszteléssel (jellemzően felhasználók bevonásával) valósítják meg. A vizsgálat során használhatnak heurisztikus módszert, kognitív bejárást, vagy tevékenységelemzést; a tesztelés pedig történhet kérdőívvel, hangos gondolkodással, vagy helyszíni megfigyeléssel. A használhatóságnál kevésbé, a hasznosság mérésénél viszont annál inkább fontos a felhasználók megfigyelése, lehetőleg „élőben”, valós körülmények között. A cikkben ismertetett vizsgálatnál az újonnan elkészült információs rendszer használhatóságát és hasznosságát egyidejűleg igyekeztek megállapítani, elsősorban a felhasználói elégedettségre koncentráva. A szakirodalomban egyetértés van abban, hogy egyetlen módszer sem ad tökéletes eredményt, és hogy az indirekt méréseket érdemes direkt módszerekkel kombinálni, mert így össze lehet vetni a felhasználók véleményét a tényleges viselkedésükkel. Ezért a kutatók a kérdőíves felmérés, valamint a helyszíni megfigyelés mellett döntöttek – ez jól bevált kombináció a korábbi tapasztalataik szerint.

Egy 18 kérdésből álló, online kitölthető űrlapot szerkesztettek, amelyben a fentiekben felsorolt jellemzőkre kérdeztek rá. Az önkénteseket arra kérték, hogy fogalmazzanak meg egy információs igényt, amely a betegellátáshoz kapcsolódik, próbáljanak választ találni rá az új e-könyvtári portált használva, majd pedig töltsék ki a kérdőívet. Eb-

ben minden egyes pontot egy 1 és 4 közötti skálán lehetett minősíteni. 4-est akkor kellett választani, ha a felhasználó teljesen egyetértett az előre megadott válasszal. Minden kérdésnél mellékeltek a szükséges fogalommagyarázatokat, és lehetőséget adtak szöveges megjegyzésekre is. Az első három kérdés a kérdőív kitöltőjére vonatkozott (életkor, beosztás, online töltött idő hetente); ezeket további három, általános jellegű kérdés követte (a rendszer rendelkezik-e az elvárt funkcionalitással?, általánosságban elégedett-e a rendszerrel?, használná-e újra a rendszert?); és az utolsó pont szintén speciális volt: ha csak egy dolgot lehetne változtatni a rendszeren, mi volna az?

A helyszíni megfigyeléshez kitalált és valós feladatokat fogalmaztattak meg a résztvevőkkel, és a saját munkahelyükön végezték el velük a tesztelést (de természetesen nem a páciensek jelenlétében). A keresési folyamatot végül is nem vették videóra, mert az ehhez szükséges közeli felvételek zavarhatták volna a tesztelő önkénteseket, továbbá időigényes a technika beállítása és az utólagos kiértékelés. Ehelyett a személyes megfigyelés mellett döntöttek, mivel ez kevésbé befolyásolja a felhasználók viselkedését. Hasonló okból a megfigyelőknek nem volt szabad reagálniuk a tesztelők esetleges megjegyzéseire vagy kérdéseire, bár ezeket is feljegyezték. Végül – a kutatást megrendelő kérésére – a tesztelőknek egy másik hasonló, fizetős adatbázisban is el kellett végezniük a keresési feladatokat és összevetniük a két rendszer jellemzőit, szintén egy 1-től 4-ig terjedő skálán (az 1 jelentette a „rosszabb”, a 4 pedig a „sokkal jobb” véleményt). A kérdőíves felméréshez és az élő teszteléshez a *Health Service* könyvtárosi hálózatán, valamint az e-egészségügy és a klinikai oktatás online fórumain toboroztak önkénteseket.

Eredmények

A kérdőíves vizsgálatra harmincan vállalkoztak (egy hiányosan kitöltött kérdőívet végül kizártak a további feldolgozásból), akik között egyaránt voltak ápolónők, szülészek, kórházi dolgozók, általános orvosok, valamint más egészségügyi szakemberek. Az eredmények összességében pozitívak voltak, de a pontszámokhoz hozzáfűzött megjegyzésekben elégedetlen hangok is megjelentek, különösen a hatékonyságra (pl. lassú válaszidő, többször is bekért jelszó), a terminológiára (nehezen érthetők egyes fogalmak és a címkék), a navigációra (túlságosan komplex rendszer, zsákutcákkal) és a relevanciára (érdektelen találatok) vonatkozóan. Mindenesetre a megkérdezettek 86%-a

igennel válaszolt arra a kérdésre, hogy használná-e újra ezt a rendszert. Az utolsó pontra pedig érdekes javaslatok jöttek: például színesebb felület, ami irányítaná a felhasználókat; online útmutató, különösen a keresőkifejezések összeállításához; egyszerű bejelentkezési pont; a találatok relevancia szerinti rendezése, és a rendszer népszerűsítése az egészségügyi dolgozók között.

A megfigyeléses tesztelésre hét önkéntes jelentkezett, de sajnos hárman rövidesen visszaléptek egyéb sürgős elfoglaltságuk miatt. A teszt során a következőkre derült fény: a felhasználókat zavarta az ismételt bejelentkezési procedúra a különböző információforrásokhoz, gondjaik voltak a keresőkérdések megfogalmazásával, és a böngésző vissza gombjával próbálták a kezdőlapra jutni, ami miatt az egyikük teljesen el is tévedt. Az utólagos megjegyzéseikből az is kiderült, hogy szerintük a navigáció nem elég átlátható, a rendszer válaszeje lassú (bár ez lehetett hálózati probléma is), a terminológia pedig nem mindenhol magától értetődő. A külalakot dicsérték, de nem tartották fontosnak. Az egyik tesztelő szerint a találati relevanciát csökkenti az, hogy túl általános fogalmakat használtak a dokumentumok tárgyszavazásához. Egy pontos bejelentkezést javasoltak, valamint összefoglalókat a dokumentumokhoz, mintakereséseket, és nagyobb hibátűrést (pl. gépelési hibák figyelmen kívül hagyása). Több esetben is összehasonlították a tesztelt rendszer képességeit az általuk kedvelt Google keresőjével, és az is kiderült, hogy nem volt nyilvánvaló mindenkinek az új rendszer célja és haszna a korábbi helyi e-könyvtárhoz képest. Megismételve a mintafeladatokat az alternatív kereskedelmi szolgáltatásban (itt a négy felhasználóból hárman kaptak releváns eredményeket), a tesztelők jobbnak minősítették a helyben fejlesztett rendszert külalak és naprakészség szempontjából, hasonlóan a hatékonyság, a relevancia és a megbízhatóság terén, és rosszabbnak a terminológiát, a navigációt és a megtanulhatóságot tekintve. A négyből hárman a második, vagyis az üzleti szolgáltatást preferálták, kiemelve annak egyszerűségét (kevesebb lépésre van szükség az eredményhez) és gyors válaszejét (ez valós körülmények között, amikor egy páciensre jellemzően 10 percnél kevesebb konzultációs idő jut, érthetően döntő faktor).

Tanulságok

Bár a kérdőíves és a megfigyeléses vizsgálat hasonló eredményeket és javaslatokat hozott, de míg

az előbbinél inkább pozitív volt az összehatás, az utóbbinál már nem annyira – az „élőben” tesztelők kritikusabbak voltak a rendszerrel szemben. Ha utólag kell kitölteniük egy kérdőívet, az emberek hajlamosak túlértékelni egy rendszer teljesítményét, mert amikor kész válaszokat kell pontozniuk, önkéntelenül azonosulnak velük; továbbá mert azt hiszik, hogy elvárnak tőlük egyfajta viselkedést, és mert a sikert vagy a kudarcot inkább a saját képességeiknek tudják be és nem a rendszer képességei következményének. Azért a kérdőíves felmérésnek is megvoltak a maga előnyei: több önkéntest lehetett így bevonni, mint a megfigyeléses módszerbe, továbbá hasznos megjegyzéseket és javaslatokat sikerült kapni a résztvevőktől. Úgyhogy érdemes ezt a módszert továbbra is használni, de kiegészítve másfajta mérésekkel, mert legalább annyira fontos azt megtudni, hogy valójában mit is csinálnak a felhasználók egy információs rendszerben, mint azt, hogy mit gondolnak róla.

A kombinált módszer előnye például akkor mutatkozott meg, amikor két tesztelő is gyorsan elvesztette az érdeklődését a rendszer iránt, mivel nem sikerült releváns információkat kinyerniük (vagyis nem bizonyult számukra hasznosnak), mégis utólag pozitívan nyilatkoztak annak használhatóságáról. Ha nem lett volna közben a megfigyelés, akkor önmagukban ezek a dicsérő megjegyzések félrevezetőek lettek volna a kutatást végzőknek. A másik szolgáltatással való összehasonlítás is hasznosnak bizonyult: a két rendszer funkcióinak összevetése lehetőséget teremtett az erősségek és a gyengeségek azonosítására. Ha a kérdőívet kitöltőknek is el kellett volna végezniük ezt az összehasonlítást, akkor lehet, hogy az ő értékelésük is kritikusabb lett volna (de ez több időt igényelt volna tőlük és valószínű, hogy akkor kevesebben vállalkoznak a feladatra).

A kutatás eredményességének két komolyabb korlátja volt: az objektívan mérhető paraméterek hiánya és a megfigyelési kísérletben résztvevők kis száma. A használhatóság és a hasznosság megítélése meglehetősen szubjektív, ezért jó lett volna a tesztalanyok véleményétől független és jól számszerűsíthető jellemzőket is mérni, de erre munkaerő és idő hiányában nem volt lehetőség. Ami pedig a tesztelők számát illeti: eredetileg 8-10 (de legfeljebb 20) résztvevőt szerettek volna a megfigyeléses vizsgálathoz megnyerni. Végül a jelentkező hétből sajnos csak négy maradt, ami statisztikai mintának nagyon kevés, még akkor is, ha ennél az időigényes módszernél amúgy is ritkán elég reprezentatív méretű az önkéntesek lét-

száma. Mindazonáltal így is sikerült értékes megfigyeléseket és hasznos visszajelzéseket kapni a résztvevőktől. Azt is fontos megjegyezni, hogy ez a kutatás csak egy „pillanatfelvétel” volt, de érdemes volna hosszabb vizsgálatokat is végezni, főleg az eredményesség, a tanulhatóság és a relevancia szempontjából. Meg lehetne például figyelni – tekintve, hogy egy klinikai döntésekhez használható információs portálról van szó –, hogy éles helyzetben (amikor a páciensek is jelen vannak), konkrét egészségügyi problémáknál hogyan alakul a fel-

használók tanulási görbéje, és hogy mennyire segíti a rendszerből kinyert információ a helyes diagnózis felállítását.

/BUCHANAN, Steven – SALAKO, Adeola: Evaluating the usability and usefulness of a digital library. = Library Review, 58. köt. 9. sz. 2009. p. 638–651. <http://www.emeraldinsight.com/10.1108/00242530910997928/>

(Drótos László)

A tudományos kommunikáció elősegítése aggregált publikációs környezetekkel

Az internet forradalmi változásokat hozott a tudományos kommunikációban. 15 éve még elsősorban a helyi egyetemi könyvtár gyűjteménye határozta meg, hogy egy kutató milyen publikációkhoz férhetett hozzá. Ami ott nem volt meg, azt hosszadalmas könyvtárközi kölcsönzéssel lehetett csak beszerezni. Manapság már azonnal elérhetők a tudományos publikációk – feltéve, hogy a könyvtárnak van elég pénze előfizetni a teljes szövegű online adatbázisokra. De minthogy nem mindenhol és nem mindenre van elegendő keret, elindult az Open Access mozgalom, aminek az az alapelve, hogy a közpénzekből finanszírozott tudományos eredmények legyenek nyilvánosan hozzáférhetők. Az elmúlt években egy sor nyílt archívum jött létre. Ezek nemcsak a kutatók közötti együttműködést könnyítik meg, hanem erősítik a tudományos világ és a társadalom közötti kapcsolatot is, hiszen így például újságírók, döntéshozók és a nagyközönség érdeklődő tagjai megismerhetik a legújabb kutatási eredményeket.

A tudományos publikálástól a tudományos kommunikációig

A kutatás során nagyon fontos a kommunikáció, és ennek csak egy része a publikálás, illetve mások publikációinak olvasása. Mielőtt egy cikk vagy könyv végleges formát ölt, a szakemberek közt rengeteg megbeszélés zajlik: nézeteket, ötleteket, véleményeket ütköztetnek – sőt esetleg a kutatási adataikat is megosztják egymással. A kiadók és az adatbázis-szolgáltatók nagyszerű eszközöket dolgoztak ki arra, hogy a szakkönyvek és szakcikkek (sőt a rájuk való hivatkozások is) könnyen és gyorsan felfedezhetőek, visszakereshetőek legyenek. A

publikálási folyamatot segítő megoldások is léteznek már, de a publikációs lánc hagyományos szereplői a szélesebb értelemben vett tudományos kommunikációt nem segítik érdemben. Tőlük függetlenül alakították ki egyes szakterületek kutatói azokat a kollaboratóriumokat (*collaboratory*), azaz együttműködési centrumokat, amelyek lehetővé teszik számukra, hogy virtuálisan szervezett kutatócsoportokban dolgozzanak együtt, közösen használjanak berendezéseket, adatbázisokat, digitális könyvtárakat vagy egyéb erőforrásokat, és természetesen közösen tudjanak így előkészíteni publikációkat is: megírni és megvitatni a kéziratokat. Jó példa minderre a komplex hálózatok vizsgálatára létrejött *Cx-Nets* (cxnets.googlepages.com) kollaboratórium, amelynek keretében amerikai, olasz és francia kutatócsoportok dolgoznak együtt. Az így megosztott információ azonban nem feltétlenül nyilvános; vagyis bár a kollaboratóriumok a modern tudományos kommunikáció minden jellemzőjével rendelkeznek, nem a nyílt hozzáférés elősegítése az elsődleges céljuk.

Értéknövelt publikációk és aggregált publikációs környezetek

Létezik egy olyan új publikációs és kommunikációs modell, amely egyesíti a hagyományos tudományos közlemények jellemzőit a kollaboratóriumok kibővített lehetőségeivel. Megteremti annak a feltételeit, hogy egy publikációt további információs objektumok (adatok, képek, hangok, vélemények stb.) egészítsenek ki, egyúttal leírja ezek kapcsolatát is a „tudományos publikációs csomagon” belül. Ez a fajta „értéknövelt” publikáció (*EP = Enhanced Publication*) nem egy statikus és lezárt dokumen-

tum, mint amilyen egy folyóiratcikk, hanem dinamikus: idővel változhat, bővíthet, megjegyzések fűzhetők hozzá, sőt egyes elemei felhasználhatók más EP-k vagy hagyományos publikációk részeként. Az ilyen komplex publikációk létrehozása technikailag nem egyszerű. Szükség van olyan eszközökre, amelyek megkönnyítik a kutatóknak az információs objektumok összekapcsolását és a köztük levő relációk leírását. Ezt a feladatot segíti az ún. aggregált publikációs környezet (*APE = Aggregated Publications Environment*), amely az *OAI-ORE (Open Archives Initiative Object Reuse and Exchange)* modellre épül.

Az OAI-ORE specifikáció webes források csoportjainak leírását és egyben ezek lehatárolását teszi lehetővé szabványos, géppel értelmezhető formában. Ilyen csoport lehetnek például: egy képgyűjtemény, egy könyv fejezetei, egy tanulmány a hozzá kapcsolódó adatsorokkal. (Részletesebben I. erről *Rumsey* és *O'Steen* cikkének bevezetőjét [1].) A modell három fő elemet különböztet meg: a forrásokat (*Aggregated Resources*), a köztük levő kapcsolatokat leíró forrástérképet (*Resource Map*) és magát az aggregációt (*Aggregation*), vagyis az így lehatárolt csoportot (1. ábra).

Az OAI-ORE modell könnyen felhasználható EP-k leírására. Az aggregáció az EP-t alkotó információs objektumokból áll össze, és ezek egymáshoz való viszonyát a forrástérkép reprezentálja. Nemcsak EP-k írhatók így le, hanem akár hagyományos publikációk is, mivel valójában ezek is több részből állnak: például *Bevezetés*, *Módszerek*, *Eredmények*, *Irodalom* – egy így felosztott, majd újra egybefoglalt tanulmánynak az az előnye, hogy egyes részei önállóan is hivatkozhatók vagy beépíthetők más dokumentumokba. Hasonló moduláris publikációkat javasolt egyébként az *Elsevier* kiadó is [3].

Az ESCAPE projekt

2008-ban a holland felsőoktatási és kutatási alap, a *SURFfoundation* tendert írt ki EP környezetek kialakítására. Ennek egyik nyertese az *ESCAPE (Enhanced Scientific Communication by Aggregated Publications Environments)* nevű projekt lett (escapesurf.wordpress.com). Az *Universiteit Twente* és a *Rijksuniversiteit Groningen* tudósai és kutatói dolgoznak együtt informatikusokkal, programozókkal és könyvtárosokkal, hogy létrehozzanak egy olyan számítógépes környezetet, amelyben könnyen lehet ilyen értéknövelt publikációkat alkotni

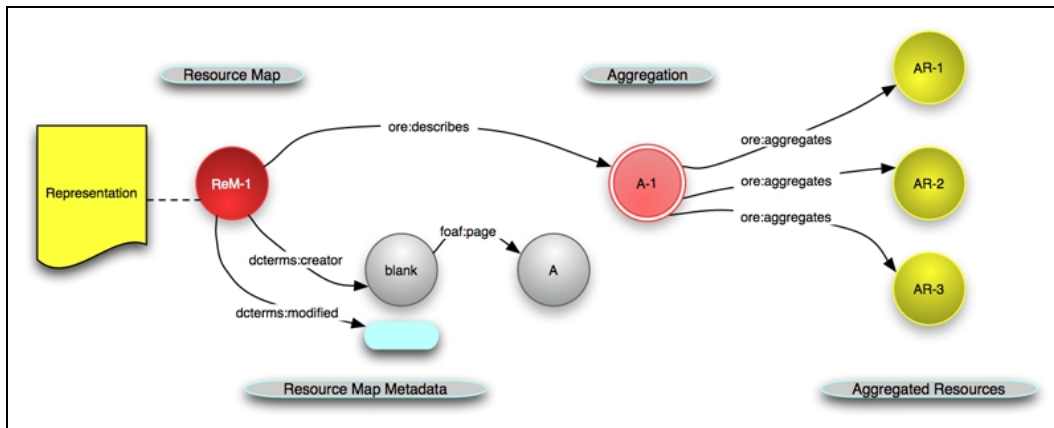
és megosztani. A világ más részein is léteznek már ehhez hasonló kezdeményezések, mint például a *Max Planck Society eSCIDoC* projektje (különbéféle információs objektumok összekapcsolása egyetlen tudományos tudástérbe), vagy a *SCOPE* nevű ausztrál fejlesztés (szerkesztőeszközök EP-k számára), illetve az *ICE-TheOREM* projekt (OAI-ORE funkciók beépítése egy tartalomkezelő rendszerbe).

A 2009 decemberéig tartó *ESCAPE* projekt keretében három kutatócsoport számára készült aggregált publikációs környezet: 1. a konfliktus, a kockázat és a biztonság témájával foglalkozó *iCRiSP* központ, 2. az ultra-nagysebességű kamera alkalmazásait fejlesztő *Brandaris 128* projekt, és 3. a közrend és a közbiztonság témájában kutató *Centrum voor Openbare Orde en Veiligheid* részére folyik ilyen *APE*-k fejlesztése. Az ezekben dolgozó kutatóknak olyan eszközökre van szükségük, amelyek nem igénylik az OAI-ORE modell és az azt megvalósító RDF nyelv ismeretét. Fontos az is, hogy ne kelljen lényegesen módosítani a már meglévő repozitóriumokon ahhoz, hogy az aggregált dokumentumokat és azok metaadatait is tárolni és keresni lehessen velük. Mivel az EP-knek – pontosabban az azok szerkezetét leíró forrástérképeknek – is saját egyedi azonosítójuk (*URI = Uniform Resource Identifier*) van, ezért ezeket ugyanúgy egyetlen információs egységként lehet kezelni, mint a szokásos (pl. egyetlen PDF állományból álló) publikációkat.

Meg kell oldani továbbá a hozzáférési jogosultságok kérdését is, hiszen az EP a kooperáció új formáinak nyit teret, de természetesen az nem lenne szerencsés, ha egy kutatócsoporton belül mindenki szabadon módosíthatná az EP-t alkotó objektumokat. Ezért az *ESCAPE* projektben három hozzáférési szintet definiáltak:

1. A szerzőknek teljes jogkörük van, szabadon bővíthetik, törölhetik vagy módosíthatják a publikációt alkotó információs objektumokat; továbbá kijelölhetnek „szerkesztőket”.
2. Ezek a szerkesztők hozzáadni tudnak (pl. címkézhetnek vagy kommentezhetnek egyes részeket), de nincs joguk módosítani vagy törölni.
3. Mindenki más csak megjeleníteni tudja az EP-t (illetve annak a forrástérképét), de nincs módosítási lehetősége.

Első lépésként létrehozták a forrástérképek archívumát, erre a célra a *Fedora 3.1* verzióját választották, leginkább azért, mert ez a repozitórium-



1. ábra Az A-1 aggregációt, amely három forrást (AR-1, AR-2, AR-3) kapcsol össze, a ReM-1 forrástérkép írja le. [2]

kezelő rendszer támogatja az RDF tripleteket (az „alany-állítvány-tárgy” típusú állításokat, amelyeket az objektumok közötti kapcsolatok leírására használnak). Ezeknek a forrástérképeknek a metaadatai ugyanúgy begyűjthetők a szokásos OAI-PMH protokollon keresztül, mint a hagyományos, egyedi digitális dokumentumok adatai. Így, ha egy felhasználó lekér egy értéknövelt publikációt a repozitóriumból, akkor egymáshoz kapcsolt objektumok hálózata jelenik meg a képernyőjén. Sőt, az OAI-ORE modellben definiált ún. *proxy* URI segítségével arra is van lehetőség, hogy egy adott információs objektumot adott kontextussal (aggregációs környezettel) együtt kérjünk le az archívumból (pl. egy bizonyos kiállítás egyik képét annak leírásával együtt).

A forrástérképek megalkotásához és módosításához az ESCAPE projekt fejlesztői egy felhasználóbarát térképszerkesztőt készítettek. Ennek segítségével az arra jogosult szakemberek relációkat hozhatnak létre és írhatnak le az objektumok között. A jövőben arra is mód lesz, hogy a felhasználók is hozzáadjanak szabadon választott leírásokat ezekhez a kapcsolatokhoz. Az EP-eket alkotó objektumok közötti relációk típusait – amennyire lehet – a már létező, elterjedt szótárakból igyekeztek összeválogatni, ezzel is megkönnyítve a későbbi együttműködést más rendszerekkel. De mivel sokféle kapcsolat lehet információs objektumok között, és ráadásul ezeket a különböző tudományterületek szakemberei különbözőképpen írják le, az ESCAPE keretében fejlesztett publikációs környezetben lehetőséget teremtettek arra, hogy a szokásos *dcterms:relation* metaadatot egy *relation-Annotation* nevű leírással is kiegészíthessék, ahol

is szabadon meg lehet adni, hogy pontosan milyen kapcsolatban van az egyik objektum a másikkal.

Tanulságok

A szemantikus web és az OAI-ORE koncepcióban komoly tudományos lehetőségek vannak. Amellett, hogy minden objektumra saját URI-val lehet hivatkozni, ezek kapcsolatai is leírhatók és ábrázolhatók egy – szintén saját azonosítóval rendelkező – forrástérképen, s ennek határai is beállíthatók. A szerzők így összekapcsolhatják azokat az elemeket, amelyeket a publikációjuk szempontjából fontosnak tartanak, az olvasók számára pedig világos lesz, hogy mik képezték az illető közlemény alapját. Vagyis az EP típusú publikálás „átláthatóbbá” teszi a tudományos kutatást, várhatóan gyorsítja is azt és közelebb hozza a tudományt a társadalomhoz.

Ezek az új fejlemények azonban új problémákat is felvetnek. Először is a forrástérképek megalkotása jóval bonyolultabb feladat, mint egyszerűen feltölteni egy publikációt egy intézményi repozitóriumba. A kutatóknak jól kell ismernie az aggregálni kívánt információs anyagokat és a köztük létrehozható kapcsolatokat. Ehhez „felülről” kell rálátnia az egész információhalmazra.

Továbbá vannak bizonyos gyakorlati – remélhetően csak átmeneti – nehézségek is: a felület nem elég felhasználóbarát, a relációk leírására használt szótárak nem teljeseek, és nem lehet még visszakeresni, hogy egy adott objektum milyen forrástérképeken szerepel. Utóbbihoz be kellene vezetni egy

olyan *ping* vagy *trackback* funkciót, amely minden alkalommal, amikor valaki felvesz egy objektumot egy EP forrástérképre, visszaküldi annak (és persze az adott objektumnak) az URI-ját annak a repozitóriumnak, amely az illető objektumot őrzi.

Vagy ott van például a névazonosítók kérdése: egy EP típusú publikációs környezetben minden objektumnál meg kell adni egy URI-t, amely az objektum készítőjét azonosítja. De sok esetben egy embernek több ilyen azonosítója is van a weben (pl. van egy oldala az egyetemi szerveren és egy másik annak a kerékpáros klubnak a honlapján, amelynek tagja); hogy mikor melyiket célszerű választani, az helyzetfüggő. Vagyis ilyenkor nagy szükség lenne az egységes digitális szerzői azonosítókra. Mindezen problémáktól függetlenül már most is vannak konkrét gyakorlati alkalmazások: A *Dataplus* projektben például a kutatók hagyományos típusú publikációkat kötnek össze kutatási adatsorokkal. Az iCRiSP központban a tudósok és más illetékes szakemberek a társadalom tagjainak a vízminőséggel kapcsolatos kockázatszlelésére és adaptív viselkedésére vonatkozó ismereteiket osztják meg egymással, az ESCAPE forrástérkép szerkesztőjét használva. Szeretnék ösztönözni ezzel is a tudományos kutatók és a politikacsinálók

közötti információcserét. Egy harmadik esetben pedig az ügyvédek tudnak létrehozni olyan EP-eket, amelyekben a hatályos joganyaghoz magyarázatok, fogalomdefiníciók, jegyzőkönyvek és konkrét jogesetek is kapcsolódhatnak.

Irodalom

- [1] RUMSEY, S. – O’STEEN, B.: OAI-ORE, PRESERV2 and Digital Preservation. = *Ariadne*, 57. sz. 2008. október 30.
<http://www.ariadne.ac.uk/issue57/rumsey-osteen/>
- [2] LAGOZE, C. – Van de SOMPEL, H.: Ore User Guide. = *Primer*, 2008. október 17.
<http://www.openarchives.org/ore/1.0/primer>
- [3] Cell Press beta, „Article of the Future”, 2009. július 20.
<http://beta.cell.com/index.php/2009/07/article-of-the-future/>

/HOGENAAR, Arjan: *Enhancing Scientific Communication through Aggregated Publications Environments*. = *Ariadne*, 61. sz. 2009. október 30.
<http://www.ariadne.ac.uk/issue61/hogenaar/>

(Drótos László)

Érdemes-e jogi és orvostudományi műveket gyűjteni az általános gyűjtőkörű könyvtárakban? Mit mondanak a használati statisztikák?

A gyűjtemények értékelését célzó projektek korábban két különböző elgondolást követtek: vagy egy-egy szakterületre koncentráltak, s a vonatkozó gyűjteményrész alapos elemzésére vállalkoztak, vagy mintavételre alapultak, s az eredményekből a teljes gyűjteményre vonatkozóan vontak le következtetéseket. Az integrált könyvtári rendszerek s egyéb korszerű technológiák megjelenése olyan eszközt adott a témával foglalkozó kutatók kezébe, amely adatok arzenálját képes szolgáltatni a könyvtári gyűjtemény tartalmi és használati szempontú értékeléséhez.

A kutató- vagy szakkönyvtárak többsége a saját gyűjtőköréhez tartozó művek mellett egyéb szakterületek alapvető munkáit is gyűjti. A cikk célja, hogy megvizsgálja, adott pénzügyi állapotok között érdemes-e a nem gyűjtőköri dokumentumok beszerzésének gyakorlatát fenntartani. A szerzők két tudományterület: a jogtudomány és a medicina

jelenlétét vizsgálják olyan könyvtárakban, ahol ezek a tudományágak nem részei a gyűjtőkörnek. A vizsgálat hét év folyamán beszerzett friss állományok kölcsönzési adatait használja fel, amelyekből éves átlagot számol. (Magyarul a hét év leforgása alatt egyszer kölcsönzött könyvet éves átlagban 1/7-szer, 0,14-szer használták.)

2006 végén a *Tudományos Könyvtárak Coloradói Szövetsége (Colorado Alliance of Research Libraries = CARL)* kezdeményezésére a monográfiabeszerzéseket összehangolandó állománygyarapítási együttműködés jött létre, amelyben a CARL-hoz tartozó könyvtárak közül nyolc vett részt hat másik („szövetségen kívüli”) könyvtárral karöltve. A projekt megvalósításához előfizettek a *Spectra Dimension* szolgáltatásra, amely lehetővé teszi az állományhasználat több szintű elemzését, valamint más könyvtárakéval való összehasonlító vizsgálatát. Annak érdekében, hogy a használat elemzéséhez megfelelő adatmennyiség álljon ren-

delkezésre, illetve a használat alakulását folyamatosan lehessen értékelni, a CARL-hoz tartozó könyvtárak feltöltötték a rendszerbe az 1999-2008 májusa közötti időszakra vonatkozó kölcsönzési adataikat. (A nem kölcsönözhető, referenzművekre vonatkozóan nem töltöttek fel adatokat a CARL könyvtárai.) Az együttműködésben részt vevő egyéb intézmények közül három csupán 1999-2005, ugyancsak három csupán 1999-2007 közötti kölcsönzési adatokkal rendelkezett.

A vizsgálatban szereplő 14 könyvtár együttes állománya hozzávetőlegesen 15 millió kötet, a hallgatók összlétszáma kb. 150 ezer, nagyjából 100 tétel jut egy hallgatóra. A kölcsönzések száma évente 1 982 000. A vizsgálatban szereplő könyvtári részállomány 1 715 000 címet tartalmaz. Egy könyvet egy évben átlagosan 0,38-szor kölcsönöztek. Az összes könyv 54,19%-át egyszer sem kölcsönözték ki a vizsgált hét évben, ez a szám könyvtárankénti bontásban 61,55%. Nagy eltérések mutatkoznak e tekintetben a nagyobb és a kisebb állománnyal bíró könyvtárak között: utóbbiak esetében az állományhasználat jóval kisebb, mint az előbbieknél. Ami a jogi és az orvostudományi könyveket illeti, a *Kongresszusi Könyvtár* osztályozási rendszere alapján készített, tematikus lebontású kölcsönzési statisztika szerint egy könyvtárban átlagosan a jogtudományi műveket kölcsönzik a legkevesebbszer (0,06), míg az orvosi könyvek a harmadik helyet foglalják el a 22 kategóriába sorolt címek között (0,39). Ha nem könyvtárakra, hanem címekre bontjuk le a kölcsönzési adatokat, azt látjuk, hogy az összes könyv közül az orvostudományi műveket kölcsönzik a második leggyakrabban! Hasonlóképpen: a jogi könyvekről mondható el a legnagyobb valószínűséggel, hogy egyszer sem kölcsönzik ki őket, míg az orvosiakra a legkevésbé jellemző ugyanez.

A jogi könyvek esetén tehát szembeötlő, hogy a könyvtárak nagy többségében kevesebbszer kölcsönzik ki őket, mint a gyűjteményben található

egyéb tematikai besorolást kapó műveket. Némileg árnyalja a képet, hogy a jogi tárgyú művek jórészt referenzállományban vannak és nem kölcsönözhetőek, ezért feltehető az erősebb olvasótermi használat is. A medicinához tartozó könyvek esetén a használat könyvtárankénti megoszlása ennél bonyolultabb képet mutat, ugyanakkor egy orvostudományi műről nagy bizonyossággal elmondható, hogy nemcsak orvosi könyvtárakban, hanem általános jellegű tudományos könyvtárakban is – legalább egyszer – kölcsönözni fogják.

A fentiekből a szerzőpáros azt a következtetést vonja le, hogy az egyes könyvtáraknak kevesebb jogi és több orvostudományi művet kellene beszerezniük, jóllehet kézenfekvőnek tűnik, hogy előbbiek esetén célszerűbb a beszerzési megbízások (approval plan) helyett a címenkénti mérlegelés lehetőségével élni. A kölcsönzési adatok helyes értelmezéséhez azonban egyéb tényezőket is érdemes számításba venni: így például az egyes könyvtárakra jellemző könyvhasználati szokásokat. Hasonlóképpen érdemes a főkategóriákon (jogtudomány, orvostudomány) túl az alkategóriákat is szemügyre venni, hogy lássuk, e két területen belül is melyek azok, amelyek a legnagyobb érdeklődésre tartanak számot. Tovább árnyalhatják a képet a digitális dokumentumok használatára vonatkozó adatok.

Mindent összevetve, korábban nem állt a könyvtárosok rendelkezésére olyan eszköz, amely az állományhasználat illetően részletetekbe menő elemzésének lehetőségét nyújtotta volna. A gyűjteményszervezés szempontjából hatalmas potenciál rejlik a korszerű technológiák ilyen célú alkalmazásaiban.

/LEVINE-CLARK, Michael – JOBE, Margaret M.: Collecting law and medical titles for general academic collections: what use statistics can tell us. = Collection Building, 28. köt. 4. sz. 2009. p. 140–145./

(Dancs Szabolcs)

Keresőgépek: valós idejű keresés

Bevezetés

A „valós idejű keresés” divatos téma lett, de hogy pontosan mit is értünk alatta, az attól függ, hogy a kifejezés melyik tagjára helyezzük a hangsúlyt. Az

egyik értelmezés szerint [1] olyan tartalmak kereséséről van szó, amelyeket gyakorlatilag azonnal közzétesznek, tehát alig telik el idő a létrehozásuk és a publikálásuk között. Vagyis például egy olyan blogbejegyzés visszakeresése, amelyet mondjuk

öt órán át írt a szerzője, nem „valós idejű”, még akkor sem, ha a *post* megjelenése után néhány másodperccel már megtalálható valamelyik keresőgéppel. Egy másik definíció [2] így határozza meg ezt a fogalmat: a megfelelő válasz megtalálása a felhasználót éppen érdeklő témában az éppen elérhető tartalmak között. Jelen cikk szerzője az első értelemben vizsgálja a *real-time search* helyzetet, vagyis hogy hogyan lehet olyan írott, képi, vagy videodokumentumokat megtalálni, amelyeket az elmúlt percekben hoztak létre és tettek fel a webre, tehát arról szólnak, ami pont most történik a világban. Ezért ebben az írásban elsősorban a mikroblog-szolgáltatások (Twitter és társai) kereséséről esik szó, és nem az olyan hírkeresőkről, mint amilyen például a Google News.

A valós idejű tartalmaknál mindig felmerül a *spam* problémája és a hitelesség kérdése. A különösen divatos vagy „forró” témákat egyesek megpróbálják meglovagolni és a saját oldalukra csalogatni a téma iránt érdeklődőket „Érdekes hír <az aktuális ügyről>”, de látogasd meg az én hogyan-legyünk-milliomosok weblapomat is” típusú üzenetekkel. Ami pedig a hitelességet illeti: a hagyományos webes műfajoknál megvan arra a lehetőség, hogy átnézzük a honlap egyéb részeit, visszafelé is elolvassuk a blogbejegyzéseket, vagy ellenőrizzük, hogy kik linkelik az adott oldalt. A valós idejű tartalmak között való tallózáskor nincsenek ilyen módszerek a közlő megbízhatóságának ellenőrzésére, legfeljebb az életrajzi adatait tudjuk megnézni, ha megadott egyáltalán valamit magáról. Vagyis bízunk kell benne, hogy az illető valóban ott van, ahol most zajlik valami, nem pedig egy olyan hamis tudósítást olvasunk, amely igazából meg sem történt.

További probléma lehet az így termelődő információ mennyisége. Ha valaki egy aktuális *breaking news* típusú eseményt követ egy blog- és mikroblog-keresővel, akkor néhány percenként százszámra kap újabb találatokat. Ehhez képest egy hagyományos, statikus weblapkereső által adott ötezer találat, amiket kényelmesen át lehet nézegetni, egészen ideális helyzetnek számít.

Valós idejű keresők

A legfontosabb mikroblog-szolgáltatónak, a *Twitternek* (*twitter.com*) természetesen van saját keresője. Sajnos a Twitter nem az elvárható gondossággal őrzi a rajta megjelenő tartalmat, rendszeresen elvesznek üzenetek, így a keresője sem

tekinthető teljes körűnek. De az alapfunkciókon kívül van néhány érdekes lehetőség az összetett keresőoldalon (*search.twitter.com/advanced*), érdemes ezért kipróbálni, ám mivel a Twitter honlapjáról nincs erre az úrlapra hivatkozás, valószínűleg kevesen használják. Pedig lehet vele földrajzi helyre és időintervallumra szűkíteni, pozitív és negatív üzeneteket keresni (az alapján, hogy :) vagy :(hangulatjelek vannak-e a szövegben), sőt kérdésekkel is (ilyenkor a ? karakter előfordulását veszi figyelembe). Hogy mennyire valós idejű a Twitter keresője, azt a szerző egy egyedi teszttüzenettel (*asd123poi456*) próbálta ki: a délelőtt 11.50-kor elküldött szöveg 25 másodpercen belül jelent meg a keresőben.

A *Sency* (*sency.com*) egy új kereső, nemrég lépett túl a béta-állapoton. A keresősor alatt két oszlopban láthatók a pillanatnyilag, illetve az utóbbi időben aktuális témák. A próbakeresés alapján feltételezhető, hogy csak a Twitter adatbázisát használja, és ez is egész gyorsan megtalálta a fenti teszttüzenetet. Nincsenek azonban további funkciói, és nincs hozzá súgó vagy „Gyakran Ismételt Kérdések” oldal. Hasznos viszont, hogy kérhetünk egy HTML kódot a keresőkérdésünkhöz, ami egy weblapba vagy blogba beépítve hírcsatornaként fog működni.

A *Collecta* (*collecta.com*) szélesebb körből merít: a Twitter mellett blogokat és azok kommentjeit, továbbá a *Jaikua*, az *Identi.ca*, a *Flickr*, a *TwitPic*, a *yFrog* tartalmát, valamint a *YouTube* és a *Ustream* videóit is leindexeli. Háromoszlopos megoldást használ, ahol is bal oldalt van a keresőmező, az éppen legfontosabb témák, a különböző opciók és egyebek; középen a találati lista, és jobb szélén egy adott találat részletesen. Az eredmények ötféle kategória szerint szűrhetők: Stories, Comments, Updates, News, Video. Ha például a friss üzenetekre vagyunk kíváncsiak, akkor az Updates opciót kell csak kipipálni, így a keresés a Twitter, a Jaiku és az *Identi.ca* tartalmára korlátozódik. A keresési eredmények megoszthatók különféle közösségi oldalakon (pl. Facebook, Delicious). A korábbi keresőkérdéseink megmaradnak addig, amíg a *Collecta* honlapján vagyunk, így rájuk kattintva megnézhetjük, hogy milyen újdonságok jelentek meg az eredeti lekérdezés óta. A rossz hír az, hogy nincs RSS csatorna és minimális a segítség a keresőkifejezések lehetséges szintaxisához.

A *OneRiot* (*oneriot.com*) szintén a valós idejű web keresőjeként hirdeti magát. Leírása szerint robotja azokat a linkeket járja végig, amelyeket az embe-

rek megosztanak egymással a Twitteren, a Digg és más közösségi helyeken, majd ezek tartalmát indexeli le másodpercek alatt. A találati listákban tehát azok a tartalmak szerepelnek, amelyek éppen az internetes kommunikáció középpontjában vannak. Így érthető, hogy az *asd123poi456* üzenetet még 45 perc múlva sem találta meg a OneRiot. Kétfajta keresési opciója van: Web és Video; és a találati lista is két szempont szerint rendezhető: Realtime és Pulse. Az első a legutóbb megosztott tartalmakat helyezi előtérbe, a második pedig azoknak ad nagyobb súlyt, amelyekre a legtöbben hivatkoztak eddig.

A *Scooper* (*scooper.com*) szintén olyan kereső, amely többféle forrást indexel: Twitter, Flickr, Digg, Delicious és egyebek. Hasonlóan két csoportra osztja a találatokat, mint a OneRiot: ám itt két oszlopban egymás mellett látjuk a „legnépszerűbb” és a „valós idejű” találatokat. Az eredmények között képek és videók is vannak, de ezeket kiszűrhetjük, ha akarjuk. A My Searches menüpont alatt a korábbi kereséseinket újra lefuttathatjuk, a Hottest Topics alatt pedig az aktuálisan legnépszerűbb témákat találjuk. A keresések szűkíthetők nagy témakörök szerint, például Technology, World Business, Sports. A Scooper az elmúlt percben közzétett tartalmakat is megtalálja, és szépen visszaadja a tesztüzenetet is időbélyegzővel, a szerzője avatárjával (az őt jelképező képecskével) és Twitter-oldalának címével együtt.

A *CrowdEye* (*crowdeye.com*) egy béta-állapotú Twitter-kereső, de a találati listája már így is nagyon informatív. A keresési eredmények időszakok szerint jeleníthetők meg, és egy grafikon mutatja, hogy mely órákban hány üzenet született az adott témában, így jól kirajzolódnak a csúcsidezőszakok. A találatok pontosítására a rendszer címkefelhőt, saját szavak beírására alkalmas szűkítő mezőt, valamint *hashtag*-alapú (# karakterrel jelölt kulcsszavakra vonatkozó) kapcsolódó kereséseket kínál fel. A középső oszlopban linklistát kapunk a keresett témában legnépszerűbb blogokról és hírforrásokról, és itt jelennek meg a releváns Twitter-üzenetek is, melyeket fontosság vagy idő szerint rendezhetünk. Érdekes specialitás, hogy minden szerző avatárja mellett egy szám is szerepel, ami az illető befolyását mutatja a Twitter-közösségen belül. Például *Stephen Fry* pontszáma 86, a *Daily Telegraph* újságé 50. A képernyő jobb szélén pedig az aktuálisan „forró” témák, és a legnépszerűbb 20 webhelyre mutató hivatkozások sorakoznak. A CrowdEye egy jól megtervezett, hatékony és könnyen használható kereső. Ha a Twitteren

kívül további valós idejű forrásokat is leindexelne, szinte verhetetlen lenne ebben a műfajban.

A fura nevű *Stinky Teddy* (*stinkyteddy.com*) egy metakereső, mely a Bing, a Yahoo!, a Videosurf, a Twitter és a Collecta keresőit kérdezi le, és megadja, hogy hány üzenet, weboldal, videó és kép található az adott témában, valamint egy *buzz-o-meter* nevű grafikon segítségével még arról is informál, hogy percenként vagy másodpercenként hány új üzenet jelenik meg a Twitteren ezzel kapcsolatban. A találatok típusok szerint vannak rendezve és szűkíthetők is e kategóriák szerint: Web, News, Video, Images, Real-time. A teszteléshez használt egyedi üzenetet a Stinky Teddy először látszólag megtalálta, de utána magának ellentmondva közölte, hogy nincs megjeleníthető találat.

Néhány további hasonló kereső: *Twazzup* (*twazzup.com*), *Twingly* (*twingly.com/microblog-search*), *Topsy* (*topsy.com*) ... és akár tucatnyit lehetne még találni, ami jól jelzi, hogy a valós idejű keresés erőteljes fejlődést mutató terület lett az elmúlt néhány évben.

Hagyományos keresőgépek

Természetesen a nagy, általános keresőgépek is igyekeznek a valós idejű tartalmakat az adatbázisukba integrálni. A Google például gond nélkül megtalálta a tesztüzenetet; a *site:twitter.com* keresőkérésre pedig 332 millió találatot adott – ezek persze tovább szűkíthetők keresőszavakkal és a Google hagyományos keresőopcióival.

E cikk írásakor, 2009 októberében – szinte egy időben – jelentette be a két nagy kereső: a *Google* és a *Bing*, hogy bekötik oldalukra a Twitter „tűzoltófecskendőjét”, vagyis a rövid üzenetek áradatát élőben követhetővé teszik. A Bing próbaverziója már működik is a *www.bing.com/twitter* címen, de használni csak akkor tudjuk, ha *United States*-re állítjuk a földrajzi helyzetünket, vagy ha eleve olyan országban vagyunk, ahol angol a hivatalos nyelv. (Rejtély, hogy miért nem tudja a Microsoft minden nemzeti verzióban egyszerre megjelentetni az új funkciókat.) A felhasználót a keresőmezőn kívül itt is egy címkefelhő fogadja a pillanatnyilag legnépszerűbb témákkal. Nem teljesen ugyanazok a *tag*-ek láthatók itt, mint amelyek a Twitter kezdőlapján vannak – a Bing láthatóan másképp súlyozza az átvett adatokat. És más a frissítés gyakorisága is: a *Windows 7* témára keresve a Bing két perccel korábban megjelent találatokat adott, mi-

közben a Twitter saját keresőjében a legfrissebb üzenet kevesebb mint 30 másodperces volt. Furcsa az is, hogy miközben a Twitter néhány perc alatt 376 új üzenetet jelenített meg, a Bing találati listája nem frissült (annak ellenére, hogy a Pause gomb nem lett lenyomva), csupán az eltelt percek száma változott az üzenetek mellett. Hogy a szolgáltatás még erősen béta-állapotban van, azt az is mutatja, hogy bár elvileg a rendszer a rövidített URL-eket (mint pl. a *bit.ly* kezdetűek) feloldja, ez a valóságban nem mindig történik meg, így a gyanútlan felhasználó esetleg spamoldalakra téved. A Bing ugyanis az üzenetek mellett a bennük megosztott legnépszerűbb linkeket is kilistázza, sőt ezek a listák tovább szűrhetők és rendezhetők olyan szempontok szerint, mint a hivatkozás óta eltelt idő, a visszahivatkozások száma és ezek íróinak fontossága. Gyermekbetegségei ellenére a Microsoft Twitter-keresője jó előrelépés, miközben a Google még nem tudott megjeleníteni egy demóval sem.

A kisebb vetélytársak közül az *Iceroocket* (*iceroocket.com*) bezzeg gyorsan átpozícionálta magát valós idejű keresővé és egész jól teljesít. Szépen megoldotta a tesztfeladatot és egy percnél nem régebbi Twitter-üzeneteket is megtalált. A Big Buzz opciót választva egyszerre tudunk keresni blogokban, a Twitteren, a video- és képmegosztó oldalakon, valamint a hírek között. A találati listáknál kérhető automatikus frissítés és a keresőkérések elmenthetők.

Következtetések

A valós idejű keresés még gyerekcipőben jár, nem alakultak ki a „kötelező” funkciók, különféle megoldásokkal folyik a kísérletezés. Azok a keresők, amelyek csak a Twitterre korlátozzák magukat, nem lesznek túl népszerűek, mert a felhasználók inkább a Twitter saját keresőjét használják – ha csak nem kínálnak olyan extra lehetőségeket, mint például a CrowdEye. A hagyományos keresők közül pedig még csak egy-kettő látott hozzá, hogy szolgáltatásába integrálja a valós idejű tartalmakat és funkcionalitást. Érdemes lesz figyelemmel kísérni a versenyt ezen a területen.

Irodalom

- [1] SULLIVAN, Danny: What Is Real Time Search? Definitions & Players. = Search Engine Land, 2009. július 9.
<http://searchengineland.com/what-is-real-time-search-definitions-players-22172>
- [2] OneRiot.com – Realtime Search for the Realtime Web. =
<http://www.oneriot.com>

/BRADLEY, Phil: Search Engines: Real-time Search. = Ariadne, 61. sz. 2009. október 30.
<http://www.ariadne.ac.uk/issue61/search-engines/>

(Drótos László)

Menekülnek a szerzők a Wikipédiáról

Egyre több ügy bizonyítja, hogy közel sincs minden rendben a népszerű online lexikon körül. A jelenségről a *The Wall Street Journal* számolt be és a gazdasági lap két videót is megjelentetett a témában. Ezek szerint szabályosan menekülnek a szerzők a *Wikipédiáról*. Csak az idei esztendő első negyedében több mint 49 000 szerző fordított hátat az online lexikonnak. Összehasonlításképpen: 2008 elején csak 4900 felhasználó döntött hasonlóképpen.

A *Wikimedia Foundation* illetékesei a hírt kénytelenek voltak megerősíteni a *The Wall Street Journal* munkatársainak, ugyanakkor közölték, hogy a jelenség főleg az angol nyelvű bejegyzések készítői körében figyelhető meg. Bármennyire is aggasztó a helyzet, a Wikimédiánál egyáltalán nincs pánik, és hallani sem akarnak a szolgáltatás megszűnéséről vagy szüneteltetéséről.

Az okokról egyelőre megoszlanak a vélemények. Sok szakértő szerint elsősorban az áll a jelenség háttérben, hogy egyre kevesebb érdekes témáról lehet írni. A mind több előírás és publikálási feltétel szintén hozzájárult a helyzet kialakulásához.

/SG.hu Hírlevél, 2009. november 25., <http://www.sg.hu/>

(SzP)