

vállalatok és magánszemélyek által egyaránt alkalmazott webnaplókat –, hanem a civil médiumok egyéb újszerű formáit (podcast, videoklip) is végigpásztázza. Nem elhanyagolható mennyiségekről van itt szó: a Technorati 54 millió blogot figyel; naponta 75 ezer új blog születik, a bloggerek vagy bloggolók 1,2 millió új bejegyzést készítenek naponta. A mennyiségen túl a hagyományos keresőmotorok még egy problémával szembesülnek. A Google és a hasonló szolgáltatások szoftverpókokat – webfűrész robotokat – küldenek ki a kiberterre, amelyek végigjárják a weblapokat, és visszahozzák a tartalmat – ezt a „hozott anyagot” indexelik aztán a keresési adatbázisban. A mennyiségi növekedéssel a webhelylátogatások közötti időszak óhatatlanul megnő. Ilyen technológiával képtelenség a blog valós idejű közegét, ezt az *instans* médiumot indexelni. A Technorati megoldása egy *pingelési* eljárás, amelyet még korábban a *Radio Userland* bloggoló-fejlesztői az új blogbejegyzés körtájékoztatására dolgoztak ki. Ha egy blogger valahol megszólalt, egy pingértésítés érkezett a szerverhez, amelyet aztán a Userland robotja felszedett, és az új bejegyzéseket indexelésre küldte. A Technorati és a nagyobb blogszolgáltatók közötti megegyezés szerint a pingelés üzemelésével az új blogbeírásokat már a közzététel utáni másodpercekben indexelik.

A Technorati eléggé hagyományos relevancia-rangsorolást alkalmaz, amelyben a fő kulcsok: a blogüzenethez mutató csatolók száma, és a tartalom belüli kulcsszó elhelyezkedése. Támaszkodik továbbá a bloggolók saját címkézésére, privát tárgyszavazására is – nem szabad persze feledni, hogy a blogkeresés messze nem egzakt tudomány. A Technorati élen járt ugyan az új műfaj menedzselésében, de a konkurencia napjainkban már erősödik. A Google például amellet, hogy megszerezte a *Blogger* szolgáltató céget, saját blogkeresőt (Blog Search) bocsátott ki, amely gyorsabbnak tűnik a Technoratinál – ez utóbbi viszont

sikeresebben birkózik meg a hatalmas mennyiségű bloghulladék (blog spam) kiszűrésével.

## ZOOMINFO

<http://www.zoominfo.com>

A *Zoominfo* erőssége a társadalmi jellegű keresés – célja az emberek, cégek, és a közöttük lévő kapcsolatrendszerek feltérképezése és megjelenítése. (Ezzel a keresőmotorral üzemel az *Amazon A9* elnevezésű keresője és a *Business Week* magazin is.) A magát összesítő keresőgépnek nevező ingyenes *Zoominfo* online forrásokból, weblapokból, hírszolgálatokból, üzleti-kereskedelmi adatbázisokból, tőzsdei állományokból szűri ki az adatokat: egy adott személynév vagy vállalati név keresőkérdésére 32 millió üzletember és 2,5 millió vállalat adatait kutatja át. A kihámozott és összesített adatsorból egy tömör és könnyen olvasható formátumú rekordot állít elő. A weben egyébként is szabadon hozzáférhető adattömegből a *Zoominfo* ingyenesen készíti a profilt: az elsősorban üzletemberek keresésére szolgáló eszköz 1 584 171 igazgatósági tagot, 323 336 vállalatvezetőt, 2 862 886 igazgatót és 128 960 jogászt tart nyilván indexében. A természetes nyelvi elemzéssel és mesterséges intelligenciával támogatott rendszer a tesztelés tanúsága szerint nem hibátlan. Tévesztés a gyakoribb neveknél érhetően többször fordul elő. A személyi profilok minősége egyenetlen, az összesítés sokszor avult. Azonosítás után egyének is kezdeményezhetik, és el is végezhetik saját profiljuk frissítését a *Zoominfo* adatbázisában. A tévesztéseket minimálisra szűkítő összetett keresés lehetősége azonban csak az előfizető használóknak jár.

WINDER, Davey: Research tools: Specialist searches. = *Information World Review*, 228. sz. 2006. október, p. 22–25. <http://www.iwr.co.uk/>

(Bánhegyi Zsolt)

## RETROKON technológia – a katalógusdigitalizálás eszköze és az internetes elérés Csehországban

A RETROKON technológia alapjait tíz éve rakták le, amikor a *Cseh Nemzeti Könyvtárban* elkezdtek a retrospektív konverziót, a papíralapú cédulakatalógus és a bibliográfiai leírások digitalizálását. Ez

az eredeti cseh technológia a nemzeti könyvtár és a *Comdat* cég együttműködésének az eredménye. Az elmúlt tíz évben ennek a technológiának a segítségével durván tízmillió, különböző cseh könyv-

tárakban található bibliográfiai és katalógustétel vált elérhetővé az interneten. A tíz év alatt a program különböző fejlődési szakaszok eredményeinek köszönhetően tökéletesedett. Az internet átviteli sebességének drámai növekedése a beszkenelt képi katalógus széles körű használatát tette lehetővé. A cseh könyvtárak katalógusainak retrospektív konverzióját azonban még nem sikerült teljes körűen finanszírozni.

A RETROKON technológia megvalósult projektjei:

- a nemzeti bibliográfia retrospektív konverziója (kb. 500 ezer tétel),
- a nemzeti könyvtár cédulakatalógusa (kb. 500 ezer tétel),
- a Szlovák Könyvtár cédulakatalógusa (kb. 300 ezer tétel),
- az Országgyűlési Könyvtár anyaga (70 ezer tétel),
- egyéb.

A RETROKON alapját négy egymásra épülő modul alkotja, amelyek önállóan és együttesen is használhatók. A technológia univerzálisan alkalmazható, tekintet nélkül az intézmény méretére és jellegére. A modulok: szkennelés és hozzáférhetővé tétel, átírás, strukturálás, revízió. A RETROKON technológia alapját a Comdat, a *TifTex* és a *ProTag* programok képezik a *RIS* és az *NRIS* programokkal együttműködve.

A technológia első fázisa a katalóguscédula szkennelése és hozzáférhetővé tétele az interneten. Ez a munkafolyamat viszonylag gyors ütemben halad, a nemzeti könyvtárban naponta 20 000 cédulát sikerült szkennelni. A szkennelési idő a katalógus tartalmától, típusától és az anyagi forrásoktól függött.

A beszkenelt cédulák képezik az internetes hozzáférés alapját a <http://katif.nkp.cz> címen, ahol a nemzeti könyvtár szkennelt katalógusa található. A könnyű és áttekinthető hozzáférés érdekében kialakították az *NRIS* rendszert, amelyen keresztül a rendszerben található bármely tetszőleges katalógus adatbázisában lehet keresni (<http://nr.is.nkp.cz>). A bevezető oldal tartalmazza a rendszer adatait, a részt vevő könyvtárakat, amelyek köre folyamatosan bővül. 2004-ben nyolc könyvtár 44 katalógusának 9 364 366 tétele volt elérhető a rendszerben.

A katalógus alapja a tételek képi megjelenítése. A használóknak a tájékozódást betűrendes faszer-

kezet teszi áttekinthetővé. A beszkenelt tételekre előjegyzés kérhető elektronikus formanyomtatvány kitöltésével. Az előjegyzés állapotáról a használó e-mail vagy webüzenetet kap. A beszkenelt képi katalógus létrehozása, összehasonlítva a teljes retrospektív katalóguskonverzióval, gyors és olcsó megoldás, átmeneti vagy végleges katalógusként egyaránt használható. A teljes katalóguskonverzió eredményeként a strukturált szöveg a nemzetközi szabványoknak és alkalmazásoknak megfelelő formátumú lesz.

Az átírás fázisában a tételek képi formátumát strukturálatlan szöveggé alakítják. Az átírás kétféleképpen történhet: automatikus karakterfelismerővel (*OCR = Optical Character Recognition*); a rossz minőségű cédulák kézi átírása a Comdat *Textif* segítségével. A strukturálásnál a szöveget a nemzeti vagy nemzetközi szabványoknak megfelelő formátumúvá alakítják az erre a célra kifejlesztett *ProTag* programmal. A tételek adatait *UNIMARC* mezőkbe és almezőkbe illesztik, ez a „maszk” könnyen átalakítható bármilyen kimeneti formátummá.

A revízió a retrospektív konverzió befejező fázisa, amely a tételek ellenőrzésével zárul. A revízió a tételek minősége szempontjából nélkülözhetetlen fázis, amelyet általában gyakorlott könyvtáros végez az eredeti cédula bibliográfiai tételei alapján. A revízió végrehajtására a *ComdatRev2003* program szolgál.

A *RIS* rendszer a retrospektív konverzió irányítására való, architektúrája kliens-szerver alapon működik hálózatos könyvtári környezetben. Az alapréteg az adatbázis, a „business logika” azokhoz a felhasználókhöz továbbítja a szolgáltatásokat, akik a rendszer kliensoldali részével rendelkeznek, vagy az interneten keresztül távoli felhasználóként csatlakoznak. A *RIS* rendszer valósítja meg és irányítja a cédulakatalógus feldolgozásának következő folyamatait: külső adatbázisból tételek keresése és átvétele a *Z39.50* protokoll alapján, retrokonverzió, rekatalogizálás.

Előjegyzés intézése: ebben a modulban a rendszer előjegyzéseket vesz fel, és nyilvántartást vezet, majd elektronikus értesítést küld a használóknak.

A tételek megjelölésével előre megadott szempontok alapján csoportosíthatjuk a rekordokat. A tételek kijelölése szabadon meghatározott szem-

pontok alapján történik (pl. nyelv, időbeli szempontok, kéziratos tétel, specifikus nyelv, specifikus írásmód, betűtípus).

A RIS rendszer fő szolgáltatása a retrospektív konverzió irányítása. A RETROKON négy fázison megy keresztül, amelyek kölcsönösen összefüggenek egymással. Az első három fázis kimenete valamilyen tételszerkezetet jelent, amely köztes terméknek számít, a következő fázisban még foglalkozni kell vele. A teljes folyamat végeredménye azonban egy megfelelő felépítésű és minőségű tétel. Az első fázisban valósul meg a szkennelés, amelynek eredménye a tétel képi megjelenítése. A második fázis az átírás, amelyben a tétel strukturálatlan szöveggé alakul. A harmadik fázisban ezt strukturált szöveggé alakítják, majd a negyedikben

a revízió során a tételek ellenőrzött strukturált szöveggé alakulnak.

A Cseh Köztársaságban a retrospektív katalógus-konverzióknak olyan minőségi, tökéletesen megbízható eszköze van, amelyet számos cseh és külföldi könyvtárban használnak. A RETROKON technológia nemcsak könyvtári, hanem más közgyűjteményi környezetben is alkalmazható.

**IBAREŠ, Miroslav: Technologie RETROKON jako komplexní nástroj pro digitalizaci a zpřístupnění katalogů, kartoték a soupisů prostřednictvím internetu. = Archivy, knihovny, muzea v digitálním světě, 2005. p. 98–104./**

(Prókai Margit)

---

## Automatizált könyvtári szolgáltatások

### Bailey Solutions

<http://www.baileysolutions.co.uk>

**Áttekintés:** A *Bailey Solutions* a Microsoft felületén működő könyvtári és információs rendszerek kezelésére szakosodott.

**Új termékek/szolgáltatások:** A PenLib a kis könyvtárak integrált rendszere; a KnowAll and Konduct a tudásmenedzsment fejlett működtetésére és integrációjára készült. A LookUp egy, a világhálón működő online katalógusrendszer, amely összeköti a hálózati OPAC-ot a hálózatalapú adminisztrációs adatlapokkal. Bárhonnan lehetővé teszi a vezetők számára a könyvtári raktár karbantartását és a katalógus keresését a munkahelyen vagy otthon. Mindegyik terméknek van OPAC rendszere a katalógus kereséséhez. A további megoldások az alapösszetevők keresésének, navigálhatóságának és használhatóságának a fejlesztésére irányulnak.

**Referenciahelyek:** A rendszer alkalmazói: Cruising Association, Field Fisher Waterhouse, Lawrence Graham és The Sugar Bureau.

### Dynix

<http://www.dynix.com>

**Áttekintés:** A *Dynix* 40 országban szolgáltat az egyetemi, szakterületi és konzorciumi könyvtáraknak. Az URSA automatizált könyvtárközi kölcsön-

zési rendszer és forrásmegosztó termék. A felhasználók a kérdéseikre egyidejű választ kapnak, a forrásmegosztást támogatja a konzorciumi és a független könyvtárak esetében is.

**Új termékek/szolgáltatások:** A Horizon Information Portal (HIP 4.0) online hozzáférést biztosít a könyvtár által birtokolt és előfizetett forrásokhoz.

**Referenciahely:** Pennsylvania Academic Library konzorcium, USA.

### EOS International

<http://www.eosintl.com>

**Áttekintés:** Az *EOS International* egy világméretű könyvtári információkezelő szoftvert és szolgáltatásokat előállító cég. Azzal, hogy Microsoft NET alapon működik, kereső- és munkafolyamat-eszközök alkalmazását teszi lehetővé. 2004-ben az EOS Web Libraryt új automatikus eszközzel, az Enterprise Reference Trackinggel frissítették, amely információval támogatja a kérdések irányítását és a gyakran feltett kérdéseket. A Web Services részegységgel a könyvtári automatizált eszközöket és az üzleti kapukat építik egybe.

**Új termékek/szolgáltatások:** Az EOS új, több adatbázist kereső rendszert fejleszt, együttműködésben a WebFeat keresőrendszer-szakértő céggel.

**Referenciahely:** KMZ Rosenman cég.