

Német webkatalógusok

Németországra nem jellemzőek sem a felvásárlások, sem a rendszer bérbevétele, ehelyett minden nagyobb portál vagy keresőszolgálat saját katalógust is épít.

- A német Yahoo!.de tételszáma 220 000; minősége kissé elmarad a Yahoo! főkatalógusától.
- A Web.de 415 000 tételével a legnagyobb. Használhatóságát korlátozza a nem eléggé differenciált osztályozási rendszer.
- A DINO online és az Allesklar tételszáma kb. 300 000, szisztematikájuk gyengélkedik.
- A német ODP 95 ezer tételével még fejlődésben van.
- A Lycos.de elve a „best of” oldalak felvétele, 55 000 tétele van.

Következtetés

A webkatalógusok továbbra is a világhálón való keresés legjobb kiindulópontjai lesznek. Nagy a koncentráció: a tételszámok gyorsulva nőnek, a versenytársak eközben fogynak; az amerikai példa szerint három-négy nagy fog talpon maradni. A nyelvi alkatalógusok minősége tartósan elmarad az angolokétól.

/HAMDORF, Kai: Wer katalogisiert das Web? = nfd Information Wissenschaft und Praxis, 52. köt. 5. sz. 2001. p. 263–270./

(Góth László)

Metaadatok könyvtári környezetben – CAMEL projekt az Oregoni Állami Egyetemen

Sokan ösztönösen vonzódnak a metaadatok világához, bennük az egyszerűbb és hatékonyabb információkeresés lehetőségét látják. Jövőbeli valós szerepüket ma még nehéz lenne megjósolni, de az biztos, hogy használatukhoz és elterjesztésükhöz körültekintéssel kell hozzáfogni. Megoldásra vár többek között a metaadatok kezelésére szolgáló eszközök jelenlegi szoftveres infrastruktúrához történő integrálása, az elektronikus források esetlegességéből (törlés, átköltötés) eredő karbantartás problémája, és a megfelelő szabványok kialakítása. Ilyen és ehhez hasonló kérdések tisztázására vállalkozott az Oregoni Állami Egyetemen (USA) 1998-ban útjára bocsátott *CAMEL (Collection And Management of Electronic Links* = elektronikus források állománya és kezelése) projekt, melynek elsődleges célja metaadatok bevonásával a könyvtár online katalógusán keresztül hozzáférést biztosítani az egyetem elektronikus forrásaihoz. Az alábbiakban a projekt során eddig felmerült társadalmi, technikai és szervezeti kihívásokról és az ezekre adott válaszokról számolunk be. (További információ a <http://osu.orst.edu/dept/library/camel/> címen található.)

A CAMEL projekt

Az egyetem oktatói és hallgatói egyre növekvő mértékben, de sokszor a vártnál kisebb eredménnyel fordultak webalapú elektronikus források-

hoz. A könyvtár nyomtatott állománya nagyobb-részt az egyes karok ajánlásait követően gyarapszik, az elektronikus dokumentumok közötti eligazodást azonban eddig semmi sem segítette. E probléma felismerése hívta életre a CAMEL projektet. A projekt a könyvtár OASIS katalógusán keresztül kívánta biztosítani az elektronikus források hatékonyabb azonosítását és elérését. Még a projekt tervezésekor eldőlt, hogy az elektronikus forrásokat leíró rekordokkal járó manuális munka csökkentése érdekében metaadatokkal fognak dolgozni.

A projekt egyedi vonásai

Az internetforrások katalogizálására szinte egyedül több kezdeményezés is vállalkozott, mint például az OCLC CORC (Cooperative Online Resource Catalog) vagy *Internet Scout* projektjei. A CAMEL projekt a könyvtárosok és szakemberek közötti hagyományos dokumentumok területén meglévő párbeszéd elektronikus forrásokra történő kiterjesztésére alapozva tűzte ki célul az internetes források feltérképezését. Ennek fényében a CAMEL projekt a következő fő vonásokban különbözik a párhuzamos kezdeményezésektől:

- A CAMEL-be kerülő elektronikus forrásokat az egyetemi karok szakemberei kutatják fel. A források a hagyományos dokumentumokhoz hasonlóan, az egyes szakterületek képviselőinek ajánlásai után kerülnek az állományba.

- Az ajánlásnak tartalmaznia kell a forrás címét, leírását és URL-hozzáférését, akárcsak egy előszerzeményezési katalógusnak a beszerzendő könyv címét, szerzőjét és tárgyszavait. Jóllehet a projekt tervezői az így nyert információt eredetileg metaadatként kívánták felhasználni, később részletezendő okokból ezek csupán támpontul szolgálhattak a könyvtárosok munkájához.
- Az ajánlást készítő szakemberek magukra vállalják az általuk megadott elektronikus források karbantartásából eredő terhek egy részét, hiszen az egyes karok elemi érdeke, hogy a számukra szükséges források hosszú távon is elérhető legyenek.
- A CAMEL projekt elkülönített internet alapú adatbázissal rendelkezik, vagyis alapvetően két részből áll: webalapú adatbázisból; OASIS rekordból. Mindez a kétlépcsős döntési mechanizmust tükrözi. Első lépésben döntés születik, hogy a könyvtár biztosítsa-e a hozzáférést az ajánlott forráshoz. Ezt követően eldöntik, hogy a forrást leíró rekord megjelenjen-e a könyvtár online katalógusában, az OASIS-ban. Ez esetben a forrásról szabvány szerinti MARC rekord készül, melynek rekordazonosítója bekerül a webes adatbázisba is. A két adatbázis következetes összehangolása teszi lehetővé valamennyi könyvtári dokumentum katalóguson keresztül történő elérését, ugyanakkor kiküszöböli a hagyományos online katalógusok elektronikus források tárolásában és karbantartásának gyakorlatában észlelhető jelenlegi hiányosságait.
- A CAMEL projekt az internetes források dinamikus természetéből adódóan csupán egy diszkrét és kezelhető halmaz feltérképezését tűzte ki célul, így az állomány fejlesztésében az Oregoni Állami Egyetemen született elektronikus publikációkra és forrásokra koncentrált.

Metaadatok alkalmazása a CAMEL projektben

A projekt beindításának egyik fő motivációja annak vizsgálata volt, hogy hogyan lehet minimalizálni az elektronikus források feldolgozásával járó katalógizálási munkálatokat. A projekt tervezőinek elképzelése szerint az ajánlásokat készítő szakemberek által kitöltött webúrlapon megjelenő információt automatikusan *Dublin Core (DC)* szabvány szerinti metaadatként, illetőleg az online katalógusba kerülő *MARC* rekord vázaként kívánták használni. Hamarosan kiderült azonban, hogy a DC szabvány pontos, következetes ismerete és használata nem várható el az egyetemi karok és szakterületek képviselőitől, így az elektronikus források jellemzése egy leegyszerűsített úrlapon történik, amely az

értékes források leglényegesebb hozzáférési pontjait megadva csupán jelzésként szolgál a könyvtárosok számára. Az úrlapon egy normál katalógushoz hasonlóan szerepeltetni kell a forrás szerzőjét, címét, URL-hozzáférését, leírását, típusát (adatbázis, kép, szöveg, levelezőlista, archívum stb.), végül az ajánló nevét.

Emberi és szervezeti kérdések

Új mechanizmusok bevezetésekor nem szabad lebecsülni az emberi tényezőket, azaz a társadalom és az egyén meggyőzését az új eszközök és módszerek alkalmazásából eredő hasznról. A projekt tervezői szerint az automatikusan generált metaadatok a CAMEL projekt gerincét képezték volna, az alábbiakban ismertetett okokból azonban meg kellett alkudni.

Az internetforrások ajánlásának elfogadtatása webúrlapon

Problémát okozott, hogy a többség a CAMEL projektet először pusztán egy újabb webkeresőnek tekintette. Jóllehet személyesen és írásban egyaránt zajlott a rábeszélés a projektben történő részvételre, igazán csak akkor vált hatásossá, mikor az elsőként ajánlott források feltűntek a könyvtár katalógusában. Munkájuk eredményét egy már ismert keresési eszközben látva, a megkeresett karok többsége fellelkesült.

A szabványos metaadatok használatának elfogadtatása

Metaadatok alkalmazásából csak abban az esetben profitálhatunk, ha következetesen használjuk őket. A projekt természetéből adódóan több szakterület (különböző karok) terminológiáját és szabványait kellett volna metaadatszinten összehangolni. Az ezzel járó nehézségeket kikerülve a CAMEL projekt a DC szabvány univerzálisan alkalmazható elemeinek elfogadtatására koncentrált, de még így is szép számmal adódtak az eltérő értelmezésből problémák.

A metaadatok közül több automatikusan is megadható, a forrásokat szignifikánsan jellemző információk nagyobb részét azonban a leírást készítő szakembernek kell definiálnia. A metaadatok következetes megadásához használt szabványok sokszor nélkülözik a leírandó forrás lényegét feltáró elemeket, vagy éppen az adott forrásra nem jellemző követelményeket támasztanak.

A karok képviselői az elektronikus forrás ajánlásához szükséges webürlap mezőit sokszor sokféleképpen értelmezték. A legnagyobb probléma a tárgyszómező kitöltéséből adódott. Az ebből származó információ egyfelől egyenetlen volt (túlrészletezett, üres, nem szignifikáns), másfelől pedig szubjektív, pénzügyi és érzelmi elemektől motivált. Jóllehet a dinamikus növekvő adatbázisnak egyre nagyobb szüksége lenne a tárgyszavak által nyújtott hozzáférésre, a fenti okokból a forrásokhoz rendelt tárgyszavakat mégsem tekinthetjük teljes értékű elérési pontnak.

A metaadatok létrehozásának elfogadtatása a könyvtárosokkal

A metaadat önmagában nem megoldás, csak eszköz. Megoldássá csak akkor válhat, ha mibenlétét megértjük, és ügyes használatát elsajátítjuk. Könyvtárainkban ma számos információforrás és hozzáférési mechanizmus létezik, azonban minden újdonsággal csínján kell bánni, mert bevezetésük könnyen visszaüthet. A CAMEL metaadatai igazodnak az OASIS katalogizálási szabályaihoz és keresési lehetőségeihez, megkímélve ezzel a könyvtárosokat és felhasználókat egy merőben új gyakorlat kialakításától. Habár ez magában hordozza azt is, hogy így a teljes szövegű elektronikus források keresési lehetőségei semmiben sem fognak különbözni a hagyományos, papír alapú dokumentumokétól, az OASIS-ban lévő CAMEL források mégis jól hozzáférhetők.

Technikai vonatkozások

A fenti emberi/szervezeti kérdések mellett természetesen nem hanyagolhatók el a projekt technikai vonatkozásai sem. A CAMEL csoport sokat foglalkozott a projekt technikai specifikációjával, remélve, hogy a kiválasztott eszközök és szoftverek a jövőben is kompatibilisek maradnak a forrásokkal és meglévő eszközeinkkel.

Számítástechnikai támogatás biztosítása

A projekt tervezési stádiumában a metaadatokkal dolgozó könyvtárak körében elterjedt, Z39.50 szabvánnyal is rendelkező *ROADS (Resource Organization and Discovery in Subject-based Services)* szoftver került a döntés előterébe. A *ROADS* üzemeltetéséhez azonban mindenképpen egy gyakorlott Unix rendszergazdára lett volna szükség, miközben használata biztonsági és adminisztrációs okok miatt az egyetem informatikai

hálózatával is ütközött volna. Mindezeket figyelembe véve a választás az egyetem szakemberei által is támogatott *Cold Fusion* rendszerre esett. A rendszer installálása és betanítása a helyi számítástechnikai szakemberek elfoglaltsága miatt meglehetősen lassan zajlott, s nem kevés időt emésztett fel a rendszer testre szabása sem.

Kompatibilitás a könyvtár információs eszközeivel és adataival

Metaadatok bevezetésekor a legnehezebb feladat a meglévő keresési eszközökkel való összehangolás. A könyvtárakban megtalálható összes forrás elsődleges hozzáférést a katalógus biztosítja, amely az információkat a könyvtárak általánosan elfogadott csereformátumában, MARC rekordokban tárolja. A pusztán számokat tekintve a MARC rekordok nem hangolhatók visszamenőlegesen a metaadatokhoz, ennek fordítva kell megtörténnie. A MARC szabvány viszont nem alkalmazható könnyen az internetforrásokra, hiszen az olyan kötelező adatmezőkkel rendelkezik, mint a kiadási dátum, kiadó vagy a formátum. Nem könnyű tehát egyértelmű választ adni olyan források esetében, amelyeket időszakosan/rendszeresen aktualizálnak, vagy egyidejűleg több formátumot is képviselnek. Az internetforrások természetüknél fogva dinamikusak, s bonyolult kapcsolatrendszer fűzi őket további forrásokhoz. Jóllehet a MARC rekord tartalmaz olyan speciális mezőket, amelyek alkalmasak lehetnének internetforrások leírására, a fentiekből adódóan ez nehéz feladat lenne.

Hiányzó szabványok

Metaadatok kezelésére a közismert erőfeszítések ellenére általánosan elfogadott szabvány jelenleg nem létezik. Az egyes érdekeltségek különböző metaadatszabványokat tekintenek magukénak. A könyvtáros társadalom érdeklődésének középpontjába — elsősorban kiterjeszthetősége miatt — a DC szabvány került, melynek további vonzerőt kölcsönöz, hogy az általa tartalmazott viszonylag kevés elem könnyen MARC rekorddá konvertálható. A DC rendkívül rugalmas szabvány, azonban kiterjeszthetőségét, egyszerűségét és az állományok/rendszerek közti átjárhatóságát nem könnyű kiaknázni, mert a szabványon belül helyenként önellentmondásba ütközhetünk. Az utóbbi években sajnos keveset fejlődött. Jelenleg valószínűtlennek tűnik, hogy a különböző felhasználói csoportok a jövőben egyöntetűen elfogadják a DC elemeinek használatát és szerkezetét.

Szoftverhátér

A metaadatok kezelését lehetővé tevő szoftverek ma meglehetősen kezdetleges szinten állnak. A legtöbbjük nem több, mint a felhasználó adatainak korrekt bevitelét szolgáló űrlapok vagy a már strukturált információt megfelelő mezőkbe tördelő programok összessége. Úttörő kezdeményezésnek számít az OCLC CORC projektjében alkalmazott rendszer, amely besorolási állományokkal is rendelkezik, és megkísérli a tartalmi analízist is.

Karbantartás

A metaadatokkal dolgozó rendszereknek megoldást kell találniuk a források esetleges törléséből, áthelyezéséből, átnevezéséből eredő kérdésekre, mivel ezek egyik legfőbb jellemzője az instabilitás. A hagyományos könyvtáru automatizálási rendszerek jelenleg még nem tudják kezelni a problémát, kérdéses tehát, hogyan lehet elérni a megváltozott forrásokat, és ki legyen ennek a felelőse. A CAMEL projekt kettős módszerrel próbálja meg a probléma feloldását. Egyfelől évente egyszer e-mailben kéri fel a rendszerben szereplő források ajánlóit, hogy ellenőrizzék az általuk ajánlott forrásokat, az észlelt változásokat pedig egy listán regisztrálják. Mindez azon a feltételezésen alapszik, hogy az ajánlóknak továbbra is érdekükben áll elektronikus forrásaik elérhetőségének biztosítása. Másfelől a CAMEL rekordban tárolt (OASIS) rekordazonosító alapján szinkronizálják a két adatbázist, biztosítva ezzel az észlelt változások következetes átvezetését.

Összegezés

A projekt e kezdeti stádiumában még korai lenne végső mérleget vonni, de már most elmondhatjuk, hogy az egyetem karai szép számmal együttműködtek az elektronikus források felderítésében. A begyűjtött források karbantartásának sikereiről még nem lehet nyilatkozni. Metaadatokat elsődlegesen a kézi munka csökkentésére alkalmaztak. Ez azonban nem hozta meg a várt eredményt, mert az ajánlók által megadott metaadatok minősége tartalmi, mennyiségi, szintaktikai szempontból nagyon egyenetlen volt.

Az elmúlt években a metaadatok kezelésére alkalmas eszközök és az ezzel kapcsolatos szabályozás nagy fejlődésen ment keresztül. Jóllehet a DC szabvány támogatottsága könyvtáros és azon kívül eső körökben egyaránt növekvőben van, jelenleg még egy szabvány sem érte el felhasználói körében azt a kritikus szintet, amely jövőben betöltött szerepüket valószínűsíthetné. Napjaink könyvtárai még kevés gyakorlattal rendelkeznek az elektronikus források hozzáféréseinek biztosításában, de minden lehetőséget meg kell ragadni a metaadatok alkalmazásával nyerhető előnyök kiaknázására.

/BANERJEE, Kyle: Challenges of using metadata in a library setting: the Collection And Management of Electronic Links (CAMEL) project at Oregon State University. = Library Collections, Acquisitions, & Technical Services, 24. köt. 2. sz. 2000. p. 217–227./

(Lengyel Monika)

Kascade: új nyílt katalógus az internetinformációk eléréséhez

Az információkeresés az interneten is a nagy számok problémája előtt áll. Sem a keresőgépek, sem az üzleti alapú katalógusok nem bírnak a feladattal, de megoldást ígér az önkéntesek tömegeinek együttműködése. A *Kascade* egy interneten alkalmazható katalógus-csereformátumra épülő nyílt rendszer.

Az információkeresés és a dokumentumszám

Az információkeresés folyamatosan megújuló módszereket igényel. A dokumentumok száma mindenféle hordozón rohamosan növekszik, ez

gyorsan elavulttá teheti a legjobb ötleteket is. 1990 körül még győzni lehetett PC-n futó szövegfeldolgozó szoftverekkel. Ám az internet minden korábbi rekordot megdönt, a weblapok száma 3 millió felett van a világon. Ez a szám évente legalább megduplázódik, és az ütem egyre gyorsul. Az új dokumentumok születését csekélyebb sebességgel kíséri a régebbiek eltávolítása. Az információkeresés nagy kihívás előtt áll.

Az információkereső szoftverek fejlődése

1964 óta vannak visszakereshető bibliográfiai adatbázisok. A STAIRS szoftver kudarcát bonyo-