

Kereső szótárak: az online keresés új típusú segédeszköze

Az online keresés térhódítása előtt az olyan eszközök, mint a tárgyszójegyzékek vagy a teauruszok, sokkal inkább az *indexelő* szakember segédeszközei voltak, mintsem a *kereső*. A nyomtatott teauruszokat ritkán tették ki a polcra a mutatókat és referátumokat tartalmazó kötetek mellé. Ennek egyik oka az volt, hogy a mutatók többnyire tartalmazták a szükséges utalásokat és az idézett források vagy a referátumok indexeléséhez felhasznált terminusokat (míg a könyvtári cédula- vagy mikrofilmkatalógusok általában csak hiányosan). A nyitott indexelő és referáló szolgáltatások nem könyvtáros használói nemigen tudták alkalmazni az indexelő szókészleteket a nyomtatott forrásokban való kereséshez.

Az online keresés térhódítása gyökeresen új helyzetet teremtett. Erőteljes igény támadt az adatbázisokban való *keresést* megkönnyítő szótárak iránt, hiszen bizonyos terminusok és kombinációik előzetes kiválasztása nélkül a számítógépes adatbázis egyetlen részét sem lehet „szemügyre venni”: a hagyományos értelemben nincs mód lapozgatásra vagy böngészésre.

A keresésért felszámított, rendszerint elég magas díjak is arra ösztönzik mind az üzemeltetőket, mind a használókat, hogy a keresések hatékonyságának növelése érdekében rendelkezésre bocsássanak, illetve felhasználjanak különféle, a lekérdezést megkönnyítő szójegyzékeket. A kereső szótárak (searching vocabularies) kifejlesztését azonban leginkább az a körülmény ösztönözte, hogy legtöbbször nem csupán az indexelő által az egyes tételekhez hozzárendelt, kötött tárgyszavak szerint, hanem a címben, a referátumban vagy a dokumentum teljes szövegében előforduló bármely szabad terminus szerint is lehet, sőt, néha kell, keresni, és a használó szívesen fogadja a segítséget ezek kiválasztásához. (A „kereső szótárak” elnevezést a szerző az „indexelő szótárakkal” való szembeállításra használja, az előbbi a kereső szakember, az utóbbi az indexelő szakember számára készül. — *A ref.*)

A kötött szókészlet rendkívül hasznos a nagy pontosságú keresésekhez, amikor azonban nincs szükség minden releváns tétel lehívására, gyakran inkább a szabadszavas kereséshez kell folyamodni a találatok növelése érdekében vagy azért, mert a kötött terminusok nem fednek bizonyos fogalmakat. Potenciálisan új szabad terminusokkal bővítik az adatbázist, valahányszor újabb tétellel gyarapítják. A szabadszavas keresés bizonyos értelemben kiszámíthatatlan, ezért a kereső szótárnak elvileg minél terjedelmesebbnek kell lennie.

A Medical Subject Headings (MeSH) szerint indexelt orvostudományi adatbázisokban való szabad kereséshez már a 70-es évek elején ún. szinoníma-teauruszt vagy használói szótárt (entry vocabulary) készítettek, hogy a

kérdéseket természetes nyelven is föl lehessen tenni a számítógépnek. A 7000 MeSH terminust mintegy 24 000 szinonímával gyarapították, és a gép automatikusan „lefordította” a szabad kifejezéseket a MeSH nyelvezetére.

A korszerű számítógépes programok csonkolással, hasonlóságok keresésével és más eljárásokkal automatikusan is képesek növelni a kereső terminusok variációinak számát, a használó rendelkezésére bocsátható kereső szókészletek jelentőségét ez azonban nem csökkenti. Egyre több, egyre változatosabb és praktikusabb típusokat fejlesztenek ki és tökéletesítenek, abból a célból, hogy a *felhasználó* ötleteket meríthesse belőlük ahhoz, milyen kereső terminusokat használjon.

A *kibővített teauruszok* és *kibővített tárgyszójegyzékek* az eredeti eszközök előnyeit megőrizve (szinonímák, hierarchikus kapcsolatok feltüntetése), a deskriptorok, illetve tárgyszavak mellett a ritkább szinonímák és nem-deszkriptorok gazdagabb választékát nyújtják, az összetett kifejezések tagjainak kiemelését permutálás segítheti, a preferált terminusok előfordulásáról gyakorisági index tájékoztathat. Az avult tárgyszavak elavulása időpontjának feltüntetése azt a célt szolgálja, hogy a használó ezeket is felhasználhassa a korábbi időszakra irányuló kereséshez.

A Library of Congress tárgyszórendszerét (LCSH) jelenleg közös vállalkozás keretében bővítik „lásd még” utalásokkal az együttműködő könyvtárak, hogy szaporítsák a hozzáférési pontokat, különösen olyan esetekben, amikor az abban szereplőknél kurrensebb kifejezések is használatosak.

A *szinonímajegyzékek* egyik példája az *Excerpta Medica* adatbázis különleges fájlja, amely 250 000 szinoníma és 200 000 preferált terminus között biztosít kétoldalú kapcsolatot.

A *tárgyszólisták* (index term listings) utalások és relációs kapcsolatok nélkül, egységes betűrendben tartalmaznak kötött tárgyszavakat, azonosítókat, összetételek tagjait, címekben, referátumokban vagy szövegben előforduló szabad terminusokat. A gépi tárolású indexek részletei megfelelő parancsra lehívhatók vagy kiírhatók. Számos adatbázis fenntartója rendszeres időközönként nyomtatott formátumban is kiad ilyen jegyzékeket.

Az *egyesített szókészletek* (merged vocabularies) két vagy több adatbázis által használt szókészlet összeolvastásával keletkeznek. Nagy előnyük, hogy elősegítik a több adatbázisban való keresést. Értékesebbek, mint alkotó részeik egyszerű összege. A *Term* adatbázis például nemrég öt társadalomtudományi adatbank tárgyszavaiból integrált új tárgyszórendszert. Sok szűkebb szakterületen ún. mikroteauruszok készülnek ezzel a módszerrel.

A különféle típusú kereső szótárak felépítése, kifinomultsága meglehetősen eltérő, az alkalmazásukkal összefüggő tapasztalatok alapján érdemes volna felmérni relatív hasznosságukat, különösen előállításuk költségeinek függvényében.

Az egyéni online használóknak és szervezeteiknek célszerű mérlegelniük, mennyire segítené elő az egyes adatbázisok használatát a megfelelő kereső szótár, és rá kell venniük az adatbázis fenntartóját, hogy készíttesse el. Az adatbázisok üzemeltetőinek ez szintén érdekük, akárcsak a meglévő szótárak állandó tökéletesítése, hogy a szolgáltatás minél jobban igazodjon a használók igényeihez.

A használói preferenciák méréséről

Amióta az információrobbanás sok könyvtárban új és költséges keresőeszközök- és módszerek alkalmazásához vezetett, a könyvtári szolgáltatások ingyenességének hagyománya megváltozott.

Az elavult manuális irodalomkutatás helyébe lépő online számítógépes bibliográfiai rendszerek lehetővé tették, hogy a használók terminál révén összeköttetést teremtsenek az olyan nagy online szolgáltató központokkal, mint pl. a Dialog, az SDC vagy a BRS. Az Egyesült Államokban több tucat műszaki, társadalomtudományi, közgazdaságtudományi adatbázisból lehet a számítógépbe táplált kulcsszavak alapján bibliográfiákat „le-hívni”. A terminálos keresés gyorsaságával, alaposágával és hatékonyságával olyan előnyöket kínál, amelyekért a korlátozott pénzügyi erőforrásokkal rendelkező könyvtárak használói a saját pénzükből is szívesen áldoznak.

Fogas kérdés, hogy a használóknak konkrétan mennyit ér meg a terminálos keresés igénybevétele. Mivel általában nem ismerik a számítástechnika nyújtotta lehetőségeket annyira, hogy választani tudjanak alternatív keresési eljárások között, a használói preferenciák nem deríthetők fel kérdőíves rendszerrel vagy interjúkkal. A számítógépes szolgáltatások fejlesztéséhez mégis egyértelműen tisztázni kell, milyen teljesítmény- és költségszintnél érdemesebb már a használónak a terminálos rendszerhez folyamodnia a manuális keresés helyett. Objektív eljárást kell tehát kidolgozni az egyes információkeresési módszerek viszonylagos hasznosságának meghatározására.

Mérési módszerek

Számos mérési módszer a manuális és a számítógépes keresés között ún. közömbösségi pont meghatározására törekedett, hogy abból következtessen a használók értékítéleteire, arra a térítésre, amely mind a könyvtár,

A kötött szókészletek összeállításainak nagyobb figyelmet kell fordítaniuk az indexelés mellett a keresési folyamat követelményeire, különösen a deskriptorok megválasztására és a képernyőn megjeleníthető tárgyszólisták, szótárrészletek megtervezésére.

/PITERNICK, A. B.: Searching vocabularies: a developing category of online search tools = Online Review, 8. köt. 5. sz. 1984. p. 441-449./

(Sz. Kiss Csaba)

mind a használó számára előnyös. Az Oregon Egyetemen végzett felmérés során például az egyetemi oktatóknak arra kellett válaszolniuk, hogy a szakterületükön végzett irodalomkutatásnál hová helyeznék el a közömbösségi pontot a két keresési alternatíva között, tekintettel olyan tényezőkre, mint idő, költség és hozzáférhetőség. Íme a kérdések:

1. Jelölje meg, mely esetben volna az Ön számára a két eljárás egyenértékű!

	terminálos keresés, amelynek költsége			
saját manuális keresés VAGY	0	50 \$	100 \$	100 \$-nál is több

2. Tételezzük fel, hogy az egyetemen van könyvtáros, aki segít az Ön specifikus témájánál, s az irodalomkutatásért Önnek kell fizetnie. Jelölje meg, mikor volna az alábbi két eljárás egyenértékű!

	könyvtáros keresés, melynek költsége			
saját keresés VAGY	0	50 \$	100 \$	100 \$-nál is több

3. Jelölje meg, mikor volna a két eljárás Önnek egyenértékű, ha az irodalomkutatást végzős hallgatóival lehetne elvégeztetni!

	végzős hallgató által lebonyolított irodalomkutatás, melynek költsége			
saját keresés VAGY	0	50 \$	100 \$	100 \$-nál is több