

val. A könyvtári teljesítménymérés jelenleg rendelkezésre álló kézikönyvei vagy csak közkönyvtárakra vonatkoznak, vagy az amerikai könyvtári viszonyokra épülnek. Az említett IFLA-munkabizottság ezt az egyoldalú helyzetet kívánja a tudományos könyvtárak javára korrigálni.

Fő törekvése az összehasonlíthatóság kimunkálása, aminek előfeltétele a mérésben használt fogalmak egyértelmű definiálása (pl. a "kölcsonzés" fogalmába beletartozik-e a hosszabítás?), a bevált módszerek összegyűjtése és szabatos leírása, a különféle struk-

túrjú könyvtárak e téren mutatkozó sajátosságainak elemzése. A munkabizottság ajánlások kidolgozásával kívánja elősegíteni, hogy a tudományos könyvtárak viszonylag könnyen megvalósítható módszerekkel összehasonlítható eredményre jussanak a teljesítményvizsgálatok terén.

/POLL, R.: Leistungsmessung in wissenschaftlichen Bibliotheken. = Zeitschrift für Bibliothekswesen und Bibliographie, 39. köt. 2. sz. 1992. p. 95–109./

(Katsányi Sándor)

A tematikus keresés szempontjainak megtervezése az online katalógusokban

A tapasztalatok azt bizonyítják, hogy a felhasználók számára igen sok problémát okoz a téma szerinti keresés megfelelő szempontjainak a kiválasztása, az igényekhez igazodó szűkítése és tágítása. Ezek a problémák két lényeges tényre világítanak rá: egyrészt, hogy kezdetben az online katalógusokban nem fektettek kellő hangsúlyt a tematikus keresésre, másrészt a témák komplexitását sem vették igazán figyelembe. A felhasználói kérdések tipologizálása megmutatja, milyen követelményeknek kell a sikeres tematikus keresés érdekében megfelelni. Az információkereső eljárások fejlesztésénél elsősorban a következő tényezőket kell megvizsgálni:

- ▶ az alkalmazott indexelő nyelvek jellegzetességei,
- ▶ az alkalmazott indexelési elvek,
- ▶ az adatbázis szerkezete,
- ▶ a visszakeresés technikai lehetőségei, a keresési mód és a keresőnyelvek.

Sikeres keresés e négy összetevő, és a közöttük fennálló kölcsönhatások átgondolt megtervezésével érhető el. Ehhez a folyamathoz kívánunk segítséget nyújtani.

A HASZNÁLAT PROBLÉMÁI

Az online katalógusok első korszakában az volt a vélemény, hogy nincs szükség tematikus elérésre, hiszen a felhasználó már ismert dokumentumokat fog keresni. Hogy ez mennyire téves álláspont, mi sem bizonyítja jobban, mint az az USA-ban végzett vizsgálat, mely szerint a keresések több mint 50%-a téma szerinti.

Ugyanakkor azonban a tematikus keresések számának jelentős növekedése nem jelenti egyszerűen a releváns találatok számának a növekedését is. Egy vizsgálat szerint a megkérdezettek 85%-a volt elégedett a kapott eredménnyel, a kérdőívek tüzetesebb vizsgálata azonban kimutatta, hogy 56%-uk semmit, vagy csak alig valamit talált meg abból, amit keresett. A használók tehát maguk sem igen tudják

megítélni, mennyire sikeres igazán egy keresés.

A felhasználói problémák főbb típusai a következők:

a) A megfelelő keresőkifejezések kiválasztása

A felhasználók általában nem találják megfelelőnek az indexelő szótár kínálat kifejezéseket, keveselik őket. Ezért aztán, a használók által kiválasztott keresőkifejezések az esetek többségében vagy túl általánosak, vagy túl specifikusak.

b) A találati arány növelésének nehézségei

A szótár nem megfelelő használatának az alacsony találati arány a következménye. A nem kielégítő keresések általában 35 és 57,5% között mozognak. Sőt, ez az arány még ennél is nagyobb, hiszen a találat sem jelenti feltétlenül azt, hogy a használó releváns információt kapott.

c) Nehézségek a találatok számának szűkítésében

Az is gyakran előfordul, hogy túl sok a találat, olyannyira, hogy már nehézkes az áttekintésük.

Főleg erre az utóbbi két pontra kell több figyelmet fordítanunk. Ahogy az információk mennyisége növekszik, egyre nő az igény a minél pontosabb és zajmentesebb online keresés iránt.

A TEMATIKUS KERESÉS FEJLŐDÉSI SZAKASZAI

Az online katalógusok első korszakának említett problémái miatt érdemes áttekinteni, milyen fejlődési szakaszokon mentek keresztül ezek a katalógusok. C. Hildreth felosztása szerint három generációt különíthetünk el.

Az első generációs katalógusok csak előre megszerkesztett kereséseket tettek lehetővé, akárcsak a kártyás katalógusok. Ezek a katalógusok általában a címrendszós keresést helyezték előtérbe, tartalmi elérés nemigen volt lehetséges. A dialógus vagy szigorú menüvezérléssel, vagy valamely parancsnyelven alapulva történt.

Csak a második generációs katalógusokban vált lehetővé a keresés kulcsszóval vagy tárgyszóval.

keresőnyelvként általában a Boole-algebra operátorai szolgáltak. Ezek már sokkal rugalmasabb katalógusok, amelyekben szabadon változtható a menüvezérlés és a parancsmód.

A harmadik generációban a felhasználói keresőprofil már sokkal jobban igazodik az igényekhez, a kérdések természetes nyelven is megfogalmazhatók. A tematikus keresésnek is sokkal szélesebbek a lehetőségei (szabad szöveges keresés stb.). Ezek a rendszerek korrekciós algoritmust is tartalmaznak, melynek segítségével a találati arány növelhető.

A TEMATIKUS KERESÉS ELEMEI

A tematikus keresés igen összetett, több elemből álló folyamat. Felfoghatjuk háromdimenziós összerakós játéknak, amelyben négy egymástól független faktor található:

- ▶ a felhasználói keresőprofil,
- ▶ az online katalógusban elérhető rekordok,
- ▶ a katalógus használói,
- ▶ a keresést megkönnyítő programrutinok.

A tematikus keresés megtervezésekor ezeket a tényezőket kell megvizsgálni, és közöttük a megfelelő egyensúlyt kialakítani. El kell dönteni, milyen szótárt használunk, milyen az indexelési mód, és melyek az indexelési elvek (mellérendelő vagy szintaktikus).

Az információkereső nyelvnek lehetővé kell tennie, hogy a felhasználó elérje az indexelő szótárt, amely az online katalógusokban általában nem túl nagy. Az elérhető dokumentumok igen sok tudományterületről származhatnak, ezért az online keresésnél sokkal nagyobb figyelmet kell fordítani a keresőkifejezések kiválasztására, megszerkesztésére és ellenőrzésére.

A helyzetet az is nehezíti, hogy az online katalógusokat – és általában a könyvtárat – nem képzett információkereső szakemberek használják. A keresőprofil tehát úgy kell megszerkeszteni, hogy a hozzá nem értő használó is eredményes tematikus keresést végezhesen. Ezért azt is meg kell vizsgálni, hogy a keresőprofil meddig segíti a felhasználót, és mikortól korlátozza a keresési lehetőségeket.

Fontos, hogy a programrutinok is segítsék a tematikus keresést. Nem mindegy például, hogy milyen a visszajelzés. A felhasználónak meg kell tudni állapítania a képernyőről, hogy a megtalált dokumentum releváns-e. A tartalomra vonatkozó információ legyen tehát világos és jól elérhető.

A FELHASZNÁLÓI KÉRDÉSTÍPUSOK

a) Egyszerű témák keresése (lexikonkérdések)

Ezek a kérdések megfogalmazhatók egyetlen szóval vagy szószervezettel, melyet az indexelő nyelv segítségével hozunk létre. Megalkotásukhoz tehát elengedhetetlen az indexelő szótárnak és kapcsolatrendszerének az ismerete. Ez a kérdéstípus meglehetősen ritka.

b) Formai adatok keresése (szerző, kiadó, kiadási év stb.)

Ez ugyan nem tematikus keresés, de az ilyen kérdések a Boole-algebra segítségével könnyen megválaszolhatók.

c) Időre és/vagy helyre vonatkozó adatok keresése

Nagyon gyakori kérdéstípus. Megfogalmazásához szükséges a megfelelő indexelemeknek (tárgyszavak, jelzetek stb.) és ezek kombinációs lehetőségeinek ismerete.

d) Összetett, a posteriori relációkra vonatkozó kérdések (pl. az irányított kapcsolatok)

Egyáltalán nem ritka kérdéstípus. Több figyelmet igényel, mivel az egyre növekvő információmennyiség elengedhetetlenné teszi ezeknek a kérdéseknek a megválaszolását. A legtöbb indexelő nyelv ad megoldást a problémára, bár ez általában erősen pre-kombinált kereséssel lehetséges csak.

A keresés megvalósításához a Boole-algebra alkalmazása szükséges. A relációk leírásához a szintaktikai eszközök kelljenek, amelyeknek az indexelő nyelvben is meg kell jelenniük.

e) Formai adatokat is tartalmazó c) vagy d) típusú kérdések

Ez a kérdéstípus gyakran előfordul a felhasználói igények között, de általában kevés indexelő eljárás nyújt lehetőséget megfelelő kielégítésére. Ezért erre a kérdéstípusra a jövőben különösen nagy figyelmet kell fordítani. Ez az a pont, ahol már nem használhatunk csak egyféle indexelő nyelvet, a különféle eljárások kombinációja vezethet csak megfelelő eredményre.

A TARTALMI MEGKÖZELÍTÉS LEHETSÉGES ÖSSZETEVŐI

Az indexelő nyelv

A hagyományos felosztás szerint beszélhetünk osztályozási rendszerekről és természetes nyelvű indexelésről.

Az osztályozási rendszerek sokoldalú használhatósága az online keresésben ismét a figyelem középpontjába helyezi őket. Főbb előnyeik a következők.

Elsősorban a témák szűkítésének és tágításának a lehetősége, amivel elősegíthető a találati arány csökkentése vagy növelése is. Ezenkívül az osztályozási rendszerek alkalmazása lehetővé teszi, hogy a nyelvi elemeket fogalmi keretbe helyezzük, elkerülve ezáltal a kétértelműséget. Segítségükkel egész fogalmi osztályokat kizárhatunk a keresésből. Az osztályozási rendszerek megkönnyítik a tematikus keresést, amely a természetes nyelvű keresőkifejezésekkel sokkal nehezebb. Ha kiderül, hogy egy bizonyos szó milyen összefüggésben szerepel az osztályozási rendszerben, a használó jobban tudja pontosítani a kérdését.

Ezek a rendszerek már a keresés során rendezet-ten adják meg a címlistát, így viszonylag nagy találati arány esetén is áttekinthetőbb a kép.

Érdekes kérdés lehet az is, hogy az osztályokat csak a jelzet megadásával, vagy az osztály megnevezésével is elérhetjük-e. Természetesen a maximális kényelmet az utóbbi variáció nyújtja. Az osztály megnevezése akár keresőnyelvként is szolgálhat, természetesen csak alapos terminológiai ellenőrzéssel, és a szükséges kiegészítésekkel. Mindazonáltal egyre gyakrabban alkalmazzák az osztályok megnevezését a természetes nyelvű indexelés alternatívájaként. Tekintet nélkül a szótár szerkezetére, a jelzetekkel való kapcsolatot minden esetben biztosítani kell.

A természetes nyelvű indexelés első pillantásra sokkal homogénebbnek tűnik, mint az osztályozási rendszerek. De el kell döntenünk, hogy egyedi tárgyszavakkal, teauruszdeszkriptorokkal vagy prekombinált tárgyszavakkal dolgozunk-e, hogy milyen mértékig alkalmazzuk a terminológiai ellenőrzést, és hogy hogyan használható a szótár. Ezt aztán rögzíteni kell, biztosítani a bármikori elérhetőséget, de nem szabad figyelmen kívül hagyni, hogy nem minden szótár alkalmas minden keresésre. Ez leginkább az indexelési elvtől függ. A természetes nyelvű keresés korlátai később jelentkeznek, amikor a találati arányt kellene növelni vagy csökkenteni. A szótár szintjén nem lehet kielégítően behatárolni a fogalmakat, vagy kizárni egyeseket. Folyamatos átmenetre lenne szükség a természetes nyelvű indexeléstől az osztályozási rendszerekhez, ez lenne az igazi segítség. A használó a nyelvi szinten kezdhetné a keresést, majd az osztályozási rendszer segítségével konkretizálhatná a kérdését.

Az indexelési elvek

Mivel a Boole-algebra segítségével a posztkoordinált keresés technikailag mindig lehetséges, így gyakran akkor is el tudunk érni minden adatot az online katalógusban, ha az nem ilyen keresésre van tervezve. Ez azonban nem mindig jelent szükségszerűen jó találati eredményeket, sőt, a prekombinált szótárak meglepő tanúsága szerint, negatívan is hathat az indexelő szótár megbízhatóságára.

A koordinált és a szintaktikus indexelést most abból a szempontból vizsgáljuk meg, hogy melyik elv miképpen befolyásolja a keresést.

a) Koordinált indexelés

A deszkriptorok egymás után következnek, anélkül, hogy bármiféle kapcsolatot határoznánk meg közöttük. Ezt az elvet főleg posztkoordinált keresésnél alkalmazzák. Egy fontos probléma, hogy milyen mértékig bontsuk szét az összetett fogalmakat. Az összetett szerkezetek egyetlen deszkriptorként szerepelnek-e (prekombináció), vagy bontsuk-e őket elemekre külön deszkriptorokká (posztkoordináció)? Mindkét eljárásnak megvannak az előnyei és a hátrányai, tehát a deszkriptorok kiválasztásánál a középútat ajánlatos választani.

Az önálló tárgyszavak, az önálló fogalmak posztkoordináltan összekapcsolhatók, de az összetett fogalmakat a Boole-algebra segítségével sem lehet tökéletesen feltárni. A koordináltan indexelt tételek nem képesek kielégítően tükrözni az összetett témákat.

b) Szintaktikus indexelés

A szintaktikus indexelés pontosabb eredményeket ad a keresésben, és a zaj is csökkenthető ily módon. A szintaktikai elemek három szinten helyezkednek el:

- ▶ Az indexelő szótár szintje (prekombinált szintaxis).
- ▶ Az indexelt tételek szintje – lehetséges-e komplex témák kialakítása a szótár elemeiből szintaktikai eszközök segítségével?
- ▶ A felhasználó szintje – tud-e a felhasználó komplex kérdéseket feltenni a rendszernek, amelyek a szótár elemeiből jönnek létre szintaktikai eszközök segítségével?

Ahhoz, hogy a posteriori relációkat tárjuk fel a szótár szintjén, prekombinált deszkriptorok kellene, ez az eljárás a prekombinált szintaxis. Magas fokú precizitást eredményez, de úgy megnöveli a szótár méretét, hogy rendkívül nehézé válik a deszkriptorok használata. A szótár szintjén ezért inkább a szemantikai (a priori) kapcsolatok feltárása a fontos.

A szintaktikai, a posteriori relációkat az indexelés szintjén kell kialakítani. Erre különféle lehetőségek vannak, pl. prepozíciók használata, deszkriptorláncok kialakítása, a deszkriptorok súlyozása, amely jelzi, hogy valamiféle kapcsolat áll fenn közöttük. Az irányított kapcsolatokat azonban csak a szerepoperátorok segítségével lehet igazán világosan feltárni.

Ha a prekombinált vagy szintaktikailag strukturált indexelemek egyes összetevőit akarjuk posztkoordinált keresésre használni, a következő problémák merülhetnek föl:

a) Van-e az egyes összetevőknek önálló szemantikai jelentésük, vagy a teljes megértéshez szükség van a prekombináltan vagy szintaktikusan indexelt tételegyüttes más elemeitől jövő szemantikai információkra is?

b) Mi módon tehetők alkalmassá a prekombináltan vagy szintaktikusan indexelt tételek a posztkoordinációra, azaz hogyan őrizhetők meg vagy állíthatók helyre a szükséges szemantikai egységek? Ha szétbontjuk a prekombinált szerkezeteket, ugyanazokkal a problémákkal kell szembenéznünk, mint az egyedi deszkriptorok esetén, és a prekombinált struktúra létrehozására kifejtett erőfeszítések is kárba vesznek. Praktikus megoldásnak tűnik a British Library PRECIS-stringeket alkalmazó keresési eljárása: a keresőkifejezés megadása után valamennyi string megjelenik, amelyben a keresett kifejezés előfordul. Így a felhasználó a prekombinált szerkezetből következtethet a dokumentum tartalmára, és a találati pontosabbak lesznek.

c) Hogyan hasznosíthatók az implicit és explicit szintaktikai eszközök a visszakeresésben?

Ezeknek a kérdéseknek a megválaszolása okozza valószínűleg a legnagyobb gondot az online katalógu-

sok tartalmi elérési elemeinek megtervezésénél. Lezögezhetjük, hogy a szintaktikus indexelés nélkülözhetetlen az összetett témák hatékony visszakereséséhez.

A különféle adatmezőkre vonatkozó keresés tervezése

A bibliográfiai tétel adatmezőinek lekérdezéséhez tudnunk kell, hogy az egyes mezők hogyan érhetőek el, és hogy egy keresésen belül kombinálhatók-e. Ehhez az adatrekord szerkezetének pontos ismerete szükséges.

A visszakeresést vizsgáló tesztek azt mutatják, hogy a kizárólag kötött szótár alapján történő keresés nem mindig eredményez jó találatokat. Ezért az ilyen kereséseket ki kell egészíteni más eljárással is (pl. szabad szöveges keresés).

Sok rendszerben alkalmaznak az egyes adatmezők eléréséhez ún. *basic indexet* (ez egy olyan összetett indextípus, amely több index kombinációjából jön létre), amelyben a szabad szöveges és a kötött szótárral való keresés együtt lehetséges. Így persze a zaj is sokkal nagyobb, hiszen ismeretes, hogy a szabad szöveges keresés mennyire zajos. A tematikus keresésnél ezért mindenekelőtt kötött szótárral dolgozzunk, csak a nagyon egyedi vagy nagyon új fogalmak esetén alkalmazzuk a szabad szöveges eljárást. Ne feledkezzünk meg az egyedi fogalmak elérése mellett a főlérendelt fogalmak elérési lehetőségének biztosításáról sem. Csak így lehet garantálni a tematikus keresés teljességét.

Az indexelt tételekhez rendelt mezőkön túl szükségünk lehet más, formai szempontokat tartalmazó mezőkre is (például a dokumentum nyelve). Ezeket is kombinálhatjuk a keresési szempontjainkkal, így tovább finomíthatók a találatok.

Az információkereső nyelv

Az információkereső nyelvet, amelyre az online katalógust alapozzuk, az indexelési elvnek megfelelően kell kiválasztani.

a) Boole-algebra

A Boole-algebra operátorai (ÉS, VAGY, NEM) az információkereső nyelvek legfontosabb összetevői között vannak. Nélkülözhetetlenek a posztkoordinált keresésnél. Használatuk azonban problémákkal jár, mivel gyakran nem eredményeznek pontos találatokat.

A Boole-algebrával történő keresésnél olyan címek is előkerülnek, amelyek jóval specifikusabb témát takarnak, mint amire a keresés irányul. Nem lehetséges csak a kívánt szavakra keresni, hiszen minden tétel találat lesz, amelyben az adott szó – akár egy hosszú deskriptorlánc egy elemeként is – előfordul.

Nincs garancia arra sem, hogy a találatokban a keresés deskriptorai egymással fogalmi összefüggésben álljanak, azaz egymástól függetlenek lehetnek, ha a dokumentumhoz két vagy több lánc tartozik.

Végül, Boole-algebrával nem lehet irányított a posteriori kapcsolatokat kifejezni (pl. a diákok értékelése a tanárok által; megjelenik a tanárok értékelése a diákok által téma is).

A Boole-operátorokat használhatjuk különböző kategóriák összekapcsolására is. Ez különösen fontos, ha a témát formai szempontokkal akarjuk behatárolni.

A Boole-algebra alkalmazása az online katalógusokban vitatott. Félő, hogy a használó könnyen elvétí a logikai kapcsolatokat. Használatára mégis elengedhetetlen komplex témák keresésénél. Jó néhány probléma elkerülhető, ha szintaktikai eszközöket alkalmazunk az indexelés szintjén.

b) A közelséget és a kontextust jelölő operátorok

A közelség és a kontextus megadásával kijelölhetjük, hogy milyen sorrendben, egymáshoz milyen közel, illetve milyen adatrekordban szerepeljenek a keresett kifejezések. Valószínűsíthető, hogy az egymáshoz közel álló kifejezések között valamilyen kapcsolat van.

c) Maszkolás (csonkolás)

Szabad szöveges keresésnél feltétlenül lehetővé kell tenni a maszkolást, hogy a morfológiailag kapcsolódó alakok is előkerüljenek. Természetesen sok lesz a nem kívánt találat is. Ez azonban csökkenthető, ha limitáljuk a csonkolást, azaz megadjuk, hogy a meghatározott karaktersort hány karakternek kell követnie. A csonkolás történhet balról, jobbra vagy középről is, így például a helyesírási nehézségek is áthidalhatók.

d) Menüvezérlés

Ahhoz, hogy az eddig említett lehetőségekkel a felhasználó igazán élni tudjon, ismernie kell ezek használatát és előnyeit. Az online katalógusnak a lehető legnagyobb mértékig segítenie kell a használatot. Ebből a célból a felhasználói interfészeket úgy alakítják, hogy azok mutassák meg a keresési kategóriákat, automatikusan gondoskodjanak a kapcsolatok létrehozásáról, illetve a maszkolásról, amennyire ez lehetséges.

A menüvezérlés segítségével a felhasználó összetettebb kérdéseket is feltehet a rendszernek. Bár ez kevésbé rugalmas, a használók mégis igen elégedettek az eredményekkel.

e) Számadatok keresése

Az információkereső nyelv fontos sajátossága, hogy képes-e számokat tartalmazó mezőket is áttekinteni. Így tovább lehet szűkíteni a keresést (pl. a kiadási évvel). Ez a fajta keresés további problémákat vet fel. Ha például nem egy konkrét évet, hanem egy időszakot keresünk, fontos, hogy ezek az adatok egy speciális numerikus adatmezőben szerepeljenek, hogy lehetséges legyen az elérésük numerikus operátorokkal.

Indextételek keresése

Különbséget kell tennünk a direkt keresés és a "böngészés" között. Mindkettőnek megvannak az előnyei és a hátrányai.

a) Direkt keresés

A keresőprofil megadása után a találat azonnal megjelenik a képernyőn. Sokkal gyorsabb, mint a böngészés, viszont sokkal kevésbé tűri a hibákat. A keresőkifejezéseket helyes sorrendben, az információkereső nyelv szintaxisának megfelelően, helyesírási hiba nélkül kell megadni. A direkt megközelítés csak akkor hasznos, ha a keresőkifejezés megfelel egy deskriptornak, vagy valamely deskriptor szinonimája. Nem deskriptorok esetén előfordulhat, hogy a rendszer behelyettesíti a deskriptort, de erről nem informálja a felhasználót, így az könnyen összevagyrodhat, mert más kifejezést lát a képernyőn, mint amit bevitt. A direkt keresés akkor válik problematikusá, ha azért nincs eredmény, mert a felhasználó valamilyen csekély szintaktikai hibát vétett.

b) Böngészés

Azt az eljárást nevezzük böngészésnek, ha a tételek megjelenítését megelőzi az indexben való keresés. Ez történhet különböző rendszerű deskriptorlistákban vagy rövidített címjegyzékekben. Ezekben az indexekben a felhasználó nyomon követheti az egymás után vagy előtt álló tételeket. Ez a keresési mód tehát hasonlít a cédulakatalógusban való kereséshez.

A böngészés lassabb, mint a direkt keresés. A felhasználó induláskor megad egy keresőkifejezést, majd miután átböngészte az indexet, választ egy újabbat. Ugyanakkor azonban ezáltal áttekintést nyer az indexelő szótárról, ami lényegesen segítheti a további keresésben.

Ha túl hosszú a lista, a felhasználó elvesztheti a türelmét, és csak felületesen olvassa át. Ezért a hierarchikus listák alkalmazását javasoljuk.

A vizsgálatok kimutatták, hogy a felhasználók többsége azonnal megáll, ha talál egy releváns kifejezést, holott lehet, hogy továbbolvasva a listát még jobbat is találna.

Az indextételek alfabetikus vagy szisztematikus elrendezésben állhatnak. Az alfabetikus deskriptorlistáknál utalni kell a szinonimákra, és tájékoztatást adni a preferált tárgyszavakról. Egyes esetekben utalhatunk hierarchikus és asszociatív kapcsolatokra is. Az alfabetikus listák gyors áttekintést adnak a szótárról, és segítenek a pontos megfogalmazásban, de hátrányuk, hogy tükrözik a fogalmak közötti kapcsolatokat.

A szisztematikus deskriptorlisták különösen hasznosak, ha meg akarjuk határozni egy kifejezés helyét a hierarchikus fogalmi szintek között, vagy ha nem kielégítő találatok esetén korrigálni akarunk. A hierarchikus fogalmi struktúráknak többféle modellje ismeretes:

- ▶ A tárgyszavak közötti hierarchikus fogalmi kapcsolatokat fa formájában ábrázolják, amely a teauruszhoz hasonlóan jelzi az alá- és fölérendelt fogalmakat.
- ▶ Egy jelzet megadása után lehetőség van az osztálynevezésekben való böngészésre.

Lehetséges az is, hogy szisztematikus keresünk a dokumentumok között. Ez olyan, mintha a polcon keresnénk. A felhasználó megad egy raktári jelzetet, amire egy szisztematikus elrendezett címlistát kap. Ez azonban feltételezi, hogy a használó ismeri a raktári jelzetet.

A megtalált indextételek megjelenítése

A megtalált dokumentum rekordja indextételeinek megjelenítése is egy lehetséges eszköze a tematikus keresésnek.

A releváns információ megtalálásakor a felhasználó előtt megjelenik az adatrekord. Ezt elolvasva eldöntheti, a cím releváns-e számára vagy sem. Ha az indextételek is láthatók, ezzel további hasznos információhoz jut, hogy mely deskriptorok vagy jelzetek alatt folytassa a keresést.

A megjelenítés formája igen változatos lehet. A további keresés megkönnyítése érdekében például bizonyos adatokat kiemelten jeleníthetünk meg. Jelzetek esetén az osztály megnevezését is meg kell adni, hogy a felhasználó könnyebben megértse, miről van szó.

A felhasználói interfész

A felhasználói interfész hozza létre a kapcsolatot ember és gép között. Úgy kell megalkotni, hogy bárki könnyedén tudjon bánni az online katalógussal, akkor is, ha nem gyakorlott a használatban. Különböző nehézségi szinteket kell kínálni, igazodva a változatos igényekhez. De az átmenet lehetőségét mindig biztosítani kell.

Sohasem szabad a legtapasztalatlanabb használókhoz igazodnia az interfész tervezésének, hiszen a felhasználónak is lehetősége van arra, hogy továbbképezze magát.

Fontos szerepük van a visszakeresésben a "help" (segítség) funkcióknak, különösen a "javaslatoknak", amelyek a különböző pontokon segítséget nyújtanak a további kereséshez. Ahhoz, hogy mindenki jól értse és használni tudja is ezeket, körültekintően kell kialakítani őket.

Javaslatok a további fejlesztésre**a) Az algoritmusok továbbfejlesztése**

Ezek olyan belső programok, amelyeket eredménytelen kereséseknél alkalmazunk. Mivel sok negatív keresés oka a helyesírási hiba, beépíthetünk például helyesírás-ellenőrző programot. Ennek a technikának az alkalmazhatósága azonban korlátozott.

Még előnyösebbek azok az algoritmusok, amelyek megkeresik a keresett szóval majdnem azonos szóalakokat.

Megkereshetők az azonos szótóhoz tartozó különböző alakok is. Ilyenkor automatikus maszkolás történik mindaddig, amíg el nem jut a program a szótóig.

Vannak algoritmusok, amelyek a szemantikai terület kiterjesztéséről gondoskodnak. Megpróbálják megkeresni a keresőkifejezéshez hasonló fogalmakat.

Egyes programok a keresőszótár ellenőrzését végzik (rendezőprogramok). Az úgynevezett "felhasználói visszajelző rutinoknál" a használó fontosságuk szerint rangsorolja az egyes keresőkifejezéseket, a számítógép pedig ennek megfelelően állapítja meg az egyes dokumentumok relevanciáját.

Ezek a módszerek természetesen elősegítik a visszakeresést, de a pontosabb találatok problémájára nem adnak megoldást.

b) A keresőszótár kiterjesztése

Sokan azt javasolják, hogy az adatrekordokat egészítsék ki a dokumentumok tartalomjegyzékéből és indexeiből vett információkkal. Így a használónak nagyobb szótár állna a rendelkezésére, tehát növekszik a találatok valószínűsége. Nagy adatmennyiség-nél azonban ez az eljárás feltétlenül problémákat okozna, főleg a terminológiai ellenőrzés hiánya miatt.

Hasonló problémák merülnek fel, ha az adatbázist különböző tárgyszavazási rendszerekkel leírt adatrekordokkal töltjük fel. Ilyenkor gyorsan megnő a dokumentumok tartalmát leíró kifejezések száma, de az inhomogenitás miatt a zajmentes keresés nem lehetséges.

c) A bevezető szótár

A már említett vizsgálatok rámutattak, milyen nehézségeik vannak a felhasználóknak a korrekt szak kifejezések megtalálásában. Sokszor inkább használják a saját kifejezéseiket, és nem törődnek a kötött szótárral.

Megoldás lehet a problémára, ha két szótárat alkalmazunk: az egyiket az indexelő szakember használja, a másikat a felhasználó. Ez a felhasználói tezausz mintegy bevezető szótárként működik. Sokkal több kifejezést tartalmaz a természetes nyelvből, így jóval terjedelmesebb, mint a kötött szótár. A kváziszinonimák mellett ez a szótár tartalmazhat eltérő írásmódú alakokat is.

A bevezető szótár terminusai utalásokkal kapcsolhatók a kötött szótárhoz. Ez a felhasználói tezausz igen hatékony eszköz, de nagyon munkaigényes. Állandó ellenőrzést és naprakészséget igényel.

Összefoglalás

Íme a főbb kritériumok, amelyek a sikeres tematikus kereséshez szükségesek:

A tematikus kereséshez az online katalógusban a dokumentumot el kell látni tartalmi leíró adatokkal, a szabad szöveges elemek csak korlátozott mértékben alkalmasak a visszakeresésre. Mind nyelvi, mind osztályozó leírásra szükség van.

A kötött szótárnak tükröznie kell a fogalmi kapcsolatokat. Ide tartoznak az alá- és fölérendelt, valamint az asszociatív fogalmakra való utalások. A további adatok (pl. a formai vagy időtényezők), amelyek a szótárelemekkel kapcsolva konkretizálják a keresést, külön adatmezőket kell, hogy képezzenek. Ahhoz, hogy összetett témákat is kereshessünk, a kötött szótárnak szintaktikus indexelési elven kell alapulnia.

A természetes nyelvű információkereső nyelvek mellett az osztályozási rendszerek struktúratükröző jelzetei is hasznosak. Az osztályozási rendszer fontos eszköze a keresés komplettebbé és pontosabbá tételének. A nyelvi és az osztályozási elemek kombinációja jobb találati arányt eredményez. A szabad szöveges keresés és a különféle kiegészítő algoritmusok csak akkor alkalmazhatók, ha egyébként nincs megfelelő találati eredmény.

A visszakereső nyelv lehetővé kell, hogy tegye a szintaktikai elemek és a számokat tartalmazó mezőkben való keresést. A felhasználónak legyen módja korrigálni a kérdését, ha a keresés túl sok vagy túl kevés dokumentumot tár fel. A szótár lehet alfabetikus vagy szisztematikus, de biztosítani kell az átmenetet a két mód között (pl. egy funkcióbillentyűvel).

A felhasználói interfész nyújtson segítséget a legjobb keresési stratégia kialakításához. Legyen lehetőség a keresőkifejezések kiválasztására, valamint a találatok bővítésére vagy csökkentésére. Mivel a menüvezérlés csak korlátozottan előnyös a visszakereséshez, a felhasználót át kell vezetni a parancsmódba.

A feladat az, hogy olyan katalógust készítsünk, amely együtt tartja a műveket, nem különíti el a kapcsolódó témákat, és könnyű keresést biztosít a felhasználó számára. (A cikkhez 73 tételből álló irodalomjegyzék tartozik. – A szerk.)

/GÖDERT, W. – HORNY, S.: The design of subject access elements in Online Public Access Catalogs. = International Classification, 17. köt. 2. sz. 1990. p. 66–76./

(Varga Katalin)