

olvasók kérdéseit, a megválaszolási módokat és a kapott eredményeket, elemezni az adott válaszok minőségét, a tájékoztató-bibliográfiai apparátus egyes részei használatának hatékonyságát. Mindezek nélkül nem lehet számítani a szakembereknek nyújtandó tájékoztató-bibliográfiai szolgáltatás megjavítására.

Az ETO alapelveiről

J. M. PERRAULT

J. M. Perrault "Towards a theory for UDC. Essays aimed at structural understanding and operational improvement." (London, Bingley, 1969. 241 p. Bibliogr. az egyes fejezetek lábjegyzeteiben.) c. munkája (Tanulmányok az ETO elméletének kialakításáról: szerkezetének megértése és felhasználóságának megjavítása céljából) a szerzőnek az ETO-ról írt 17 tanulmányát tartalmazza, négy nagy fejezetbe csoportosítva, a tiszta elméleti megalapozástól a gyakorlati kérdések felé haladás rendjében. A négy fejezet címe: 1. Általános elméleti alapelvek. 2. Az ETO szerkezete. 3. A jelzetelés kérdései. 4. Az ETO és az újraosztályozás. Ezek közül jelenleg csak az 1. fejezetnek, az elméleti alapvetésnek ismertetésére szorítkozhatunk.

DOKUMENTÁCIÓS RELEVANCIA ÉS STRUKTURÁLIS HIERARCHIA

Minden visszakeresési rendszer alapvető problémája a relevancia mértéke. A relevancia értékítélet kérdése. Lehet valami releváns egy adott dokumentum indexelésekor az indexelő számára, ami a keresőnek nem az, és megfordítva. Az információelméletben ezt a megnevezést nevezik a jnak, de fokozódhat ez c s endé is, mikor a visszahívási arány egyenlő a nullával. A várható kérdés nem látható előre, így a dokumentum-szurrogátum készítője (az indexelő, a szakozó) ehelyett a dokumentumnak – maga konkrétságában – az adottságait rögzíti. A keresésnél kvázi-egybeesések vannak az indexelésekor adott fogalmakkal. A gépi visszakeresés követelménye, hogy a formális megfelelések egyenértékűek legyenek a relevanciával, és hogy az egész eljárás mechanikusan elvégezhető legyen. Hogyan érhető ez el? Minden fogalom, amely megadja a dokumentum tartalmát (és pl. A, B, C, D kifejezésekkel rögzíthető) speciei genusának tekinthető: $A = A_1, A_2, \dots, A_n$. Az adott példában $A =$

miniatűr kutyafélék, A_1 = kínai pincsik, tehát az A, B, C, D fogalmi tartalom releváns az A_1 -re is. Sajnos ez a fajta egybeesés csak az explicit relevanciának nevezhető esetekre megoldás. Komplexebb esetekben ez nem ilyen egyszerű, azonban annyit ki lehet e példából szűrni, hogy a relevancia egyik jellegzetessége, hogy megítélése hierarchikus értékítélet. Ennek a felismerésnek a birtokában a szerző nem tud egyet érteni azokkal (és sokan vannak), akik azt állítják, hogy az "osztályozás" nem mindig hierarchikus. Az "osztályozás" kifejezés csakis a hierarchikus alapon történő tartalmi feltáráshoz (dokumentum-szurrogátum készítésre) alkalmazható joggal. Kifejezetten meg kell különböztetni a (hierarchikus) osztályozást az "indexelés"-nek nevezett eljárásoktól. Maga az indexelés kétféle típusu lehet aszerint, hogy szókészlete ellenőrzött vagy szabad. Ha szókészlete ellenőrzetlen, akkor leírja ("de-skribálja") a dokumentumot a benne foglalt szavak segítségével, ha ellenőrzött, akkor megadja ("a-skribálja") a jellemző fogalmakat. Tehát az indexelés mindkét esetben a dokumentumelemek (-)skripciója. Ezzel szemben az osztályozás pozíciók előírása (pre-skripciója) valamely szisztematikus fogalmi rendszerben, a dokumentumokat pedig in-skribálja ezekbe a pozíciókba. Így az osztályozás is (-)skripció, de nem a dokumentumelemeké, hanem az egész dokumentumé az egésznek tekintett corpusba.

Az osztályozás szisztematikus fogalmi szervezete bizonyos mértékig jellemző a tézauruszok tárgyszavaira is. Viszont egy hierarchikus osztályozási rendszerben sem áll minden osztály hierarchikus kapcsolatban az összes többi osztállyal, de implicit szisztematikus kapcsolatok mégis vannak köztük. A fogalmak közti hierarchia a strukturált jelzetekben kifejezve válik alkalmassá a számítógépbe való betáplálásra és visszakeresésre. A gépi indexelés hívei azt állítják, hogy az osztályozás – tehát valamiféle dokumentum-szurrogátum készítése – alkalmatlan a gépi visszakeresésre, csupán a dokumentum szerzője gondolatainak saját szavaival való kifejezése és ezeknek a dokumentum teljes szövegéből való kivonása alkalmas erre. Az eddigi eredmények ezt még nem igazolták kellőképpen. A szerző szilárdan kitart meggyőződése mellett, hogy a fogalmi feltáráson alapuló bibliográfiának egyetlen módja az osztályozás. A kérdés ez: lehetséges-e számítógéppel előállítani "tökéletes" preskripciót? Vagy ez is izlés dolga, mint a kézi osztályozások esetében? A feladat tehát az osztályozási rendszerek (kézi vagy gépi) osztályozása abból a szempontból, mennyire felelnek meg saját alapelveiknek szerkezetükben, és mennyire képesek biztosítani a relevanciát.

A FOGALMI ELEMELK FELTÜNTETÉSI SORRENDJÉNEK ELVI ALAPJAI, SZERKEZETE ÉS FUNKCIÓJA

A fogalmi elemek feltüntetésének sorrendje (citation order, az adekvát magyar kifejezés hiányában a továbbiakban CO-nak fogjuk nevezni), egy információ-tárolási és -visszakeresési rendszerben kereshető (megtalálható) szurrogátumok megteremtésére irányuló szabály, amely előírja a kifejezések mesterséges nyelvi formába való rendezését. Kaiser kategóriái és CO-ja: konkrét + folyamat. Ranganatban minden komplex mesterséges nyelvi kifejezésformának alkotóelemeit, a kifejezéseket (ter minusokat) elemezve valamennyiüket besorolja öt alapvető kategóriáinak (személy, anyag, energia, tér és idő) egyikébe

és előírja, hogy a kifejezéseket ebben a sorrendben kell kombinálni. Vickery az alapvető kategóriákat tízre növelte: dolog, fajta, rész, tulajdonság, anyag, folyamat, tevékenység, ható tényező, tér és idő; feltüntetésüket ebben a sorrendben írja elő. Mi ezeknek a szabályoknak elméleti és tényszerű alapja és alkalmazásuknak hatása? Ez a kérdés nemcsak a szisztematikus (jelzetelő), hanem a verbális tájékoztatási nyelvek szempontjából is egyaránt fontos. 1. Az első alapvető feltételezés a CO mélyén a prekoordináció, vagy másként az artikuláció, a komplexumok megléte. Ennek ellentéte Taube uniterm-javaslatára, melynek pl. egyik át nem gondolt következménye a tárgyszavazásban az ilyesfajta eljárás: egy dokumentumnak, amely egy térképalbum Európa gazdasági helyzetéről, három tárgya van: TÉRKÉPEK, EURÓPA, GAZDASÁGI VISZONYOK. Tény, hogy a dokumentum az európai gazdasági viszonyokról szól, de nem a térképekről. Taube hatásának másik bizonyítéka, hogy a tézauruszokban kerülnek a komplex kifejezéseket a potenciális tévedések elkerülése végett. Ezzel szemben: a jelentés nem az izolált kifejezésekben, hanem a kifejezések kapcsolatában, egyszóval: kifejezés-komplexumokban adott. 2. A CO második alapfeltételezése, hogy ha a jelentés kifejezés-komplexumokban adott és ha kommunikációs eszközeink egy dimenziósak, akkor a komplexumot kitevő kifejezések sorrendje határozza meg kapcsolatukat. Ranganathan példája: "az osztályozás története" és "történelem osztályozása". 3. A CO határozza meg a gyűjtemények felállításánál, hogy milyen szempontok szerint kerülnek össze a művek. Pl. a Library of Congress osztályozási rendszere(LC) szerint az irodalom-szakban a CO: nyelv: korszak: szerző: műfaj. Emögött az a feltételezés áll, hogy a tudós érdeke, hogy egy szerző műveit és a róla szóló irodalmat együtt találja meg, ugyanakkor mellette ugyanazon korszak irodalmát. A Tizedes Osztályozás irodalom szakcsoportjának CO-ja: nyelv: műfaj: szerző. A művelődni kívánó olvasó (általános, tehát nem tudós) érdeke, hogy egy nyelv irodalmának azonos műfajait együtt találja meg. Természetesen a feltételezések érvényessége nyitott kérdés és Taube kísérlete főként az egyetemességnek ebből a hiányából ered.

A komplex jelentés, a komplexumok sorrendje és a variabilitás alapvető feltételezéseiből kiszűrhető az a megfigyelés, hogy a CO alapelve a csökkenő konkrétság (ld. fentebb Kaiser, Ranganathan és Vickery kategóriáit), míg a besorolás alapelve: az általános megelőzi a speciálisat. Másik érdekes megfigyelés a CO és a besorolási rend invertált viszonyáról az, hogy mindegyik kategória a sorban azáltal konkrétabb, mint a másik, hogy osztható a másik által. Negatíven megfogalmazva: ha egy P kategória kifejezései nem képesek felosztani Q kategória kifejezéseit, akkor Q jön előbb a CO szerint. Egyszerű mátrix segítségével is bemutatja ezt egy X osztály A, B, C, D kifejezésein, ahol a konkrétság mértéke A-tól D-ig növekvő s így bizonyítottan veszi a tételt, hogy "az általános megelőzi a speciálisat" belsorolási rend nincs konfliktusban a CO-elvel: "a konkrét (speciális) megelőzi az absztraktot (általánost). Minden táblázatot (bármilyen rendszernél) a besorolási renddel azonos formában szerkeszteni meg.

A CO tehát alapvető szabály minden információ-visszakeresési rendszerben, alapja valamennyi vetületés vagy analitikus-szintetikus osztályozási rendszernek, implicite olyan osztályozási rendszerekben is felfedezhető, amelyek nem rá épülnek. Az USA-ban is, ahol pedig a szakkönyvtárak és a tájékoztató központok hagyományosan ragaszkodnak Taube elméletéhez és gyakorlatához, erősen megnőtt az érdeklődés az osztályozási rendszerek iránt. Taube és köve-

tői a fogalmaknak nem jelzet-megfelelőit, hanem verbális megfelelőit használják. A tárgyszavazást azonban, amely már régi verbális keresési stratégia, éppen annyira támadják, mint a jelzetelő osztályozásokat, mivel implicite a tárgyszavak is alkalmaznak CO-t. De a szerző számára a legfontosabb az a felismerés, hogy a CO szabályozza egy tézauruszban adott kifejezésekből alkotott komplex tárgyszavakat is. Meg kell határozni a kategóriákat, amelyek egy speciális terület komplex kifejezéseit alkothatják, különösen olyan jelentésváltozásoknál, ahol a jelentés a sorrendtől függ. Pl. a "bibliográfia" kifejezés jelenthet egy problémát ("átfogó bibliográfia szükségessége az ázsiai biológiai kutatásokról" címben pl.) vagy egy formát ("zárvatermő növények irodalmának bibliográfiája"). Szimbólumokat kell bevezetni a komplex kifejezés valamennyi tagjának kategóriáit illetően. De maga a szókészlet is tézaurusz-szerűen, azaz hierarchikusban strukturált. Index mutatja a felhasznált kifejezések pozícióját és szinonimáit. A komplex tárgyszavak között lásd még reciprok hivatkozások állapítandók meg. A strukturálás következménye az indexelő nyelvben, hogy a tárgyszót követheti pl. egy formát kifejező terminus, egy helyet kifejező terminus vagy mindkettő; tematikusan elosztó kifejezés, ez és egy formát kifejező terminus, vagy egy helyet kifejező terminus, vagy mindhárom. Egyszóval a jelzetelő osztályozási rendszerekre alkalmazott mátrix-elemzés világosan feltalálható a verbális tájékoztatási nyelveknél is. A hierarchikus tézauruszok (amelyek alkalmazzák a komplex kifejezések permutációját és a CO-t) állnak a legközelebb az osztályozási rendszerekhez és használják ki leginkább azok előnyeit – verbális rendszerben.

REFLEXIÓK AZ ÁLTALÁNOS ÉS SPECIÁLIS KAPCSOLATRÓL A VERBÁLIS KERESÉSI STRATÉGIÁKBAN

A "lásd még" hivatkozások jól ismert szabálya a tárgyszórendszerben: hivatkozz egy általános tárgyszóról egy kapcsolt, de speciálisabb tárgyszóra; vagy a specificitás azonos szintjén álló kapcsolt tárgyszóra; de ne hivatkozz egy kapcsolt, de általánosabb tárgyszóra. A Library of Congress tárgyszó-rendszere (LCSH) és a Sears' list of subject headings ezt a szabályt alkalmazza, bár egyes – tegyük hozzá: üdvös – kivételek is vannak ez alól. Pl. a Görög irodalomról és a Latin irodalomról hasznos a hivatkozás a Klasszikus irodalomra, bár a szabálynak ellentmond. Az információ-visszakeresés terén egy másik jól ismert szabály, hogy "vedd ki az A, B, C-nek vagy A, B, D-nek, vagy A, B, E-nek kódolt dokumentumokat, ha a keresés megfogalmazása A, B; de ne vedd ki ugyanebben a keresésben az A, C-nek, vagy B, C-nek, vagy A, D-nek kódolt dokumentumokat. Ennek a szabálynak alapja hasonló a "lásd még" hivatkozások "általánosról a speciálisra" irányulásához. A tézauruszok BT-i (Broader Term = tágabb kifejezés), RT-i (Related Term = kapcsolt kifejezés) és NT-i (Narrower Term = szűkebb kifejezés) megfelelnek a "lásd még" hivatkozásoknak, de a nem-hagyományos "speciális ld még általános" hivatkozásnak is. A posztkoordinált rendszerekben való keresésnél pedig tudatosítani kellene, hogy a nem kívánt specifikáció (A, B, C

mint válasz az A, B-re) nem kevésbé veszélyes, mint ellentéte, a semmitmondó általánosítás. Példa ez utóbbira: egy farkas kutyaféle ugyan, de az emlős mint olyan nem kutyaféle. Az érvelés kétirányú: a következő állítás "a farkas egy állat" nem fordítható vissza így: "az állat egy farkas". Az igazi probléma ott van, hogy nem minden további kifejezés specifikáció. Ha a kérdés így szól: "bibliográfia (A) háromszögszárnyu (B) repülőgépekről (C)", akkor a B a C specifikációja, de az A nem. Így az A, B, C-nél az A-t nem lehet kiküszöbölni (amennyiben nem kívánt formájú dokumentumot akarunk megtalálni), de a B-t vagy a C-t kiküszöbölhetjük, ha az A, B, C kérdésre találunk néhány vagy elég dokumentumot, amelyeket A, B-nek vagy A, C-nek kódoltunk.

A SZURROGÁTUMOK TAGOLÁSÁNAK ISMERETELMÉLETI ALAPJAI

Az ismeretelmélet minden, reláción alapuló rendszernél implicite jelen van. A jelzést (indikációt) információvá a szemantikai elemeknek szintaktikai elemek segítségével történő átalakítása teszi. Ősrégi ismeretelméleti tétel, hogy "az igazság az ítélet által található meg". Az ítélet legelemibb jelentésében az az értelmi tevékenység, amellyel kapcsolatot létesítünk két (vagy több) fogalom vagy létező között. Ami az ítéletet lehetővé teszi, az a hagyományos filozófiában a kopula (a "van", a létige). A Boole-féle logikai módszerekkel meg is kísérlük valamennyi viszonylat redukálását a "van"-ra és a "nincs"-re. De milyen szegényes a dokumentáló Boole-féle logikája a szintakszis valódi gazdagságához képest! Heidegger mutatja ki többek között a létige (van) sokféle jelentését. De a szintakszis funkciója, hogy a strukturáltat egybefűzi, így a következő lépés a tagolás (artikuláció), amely az ismeretelmélethez kapcsolja a nyelvészeti elemzést. Az artikuláció teszi lehetővé a beszédet, amely az elszigetelt állításokkal szemben eljuttat minket a dokumentumokban foglalt igazsághoz (vagy tévedéshez). Az igaz tudáshoz (scientia) való hozzájárulásként egy dokumentumnak valami újat és igazat kell tartalmaznia. A dokumentumszurrogátumoknak – amelyek visszakereshetővé teszik a dokumentumot, amennyiben releváns a kérdésre – artikuláltaknak kell lenniük, azaz a szurrogátumoknak is van valamiféle szintaksziszuk, amelyet nem szabad túlzottan leegyszerűsíteni, azaz megállni a szurrogátumok in-dikativ tárolásánál (amelyek a beszédben a tiszta fogalmaknak felelnének meg). A tiszta fogalmakhoz olyan kvázi-sziktakszist kell még adni, amely a szurrogátumot nem beszéddé artikulálja, hanem minősíti a tiszta fogalmakat. Egy olyan egyszerű mondatban, mint "ha János meghalt, akkor Mária bánatos lesz", a négy tiszta fogalom – János, meghalt, Mária, szomorú – artikuláció nélkül pl. olyan értelmezéshez is vezethet, hogy: "ha Mária bánatos, akkor János meghal". A kvázi-szintaktikus elemek (a szerepjelölők = roles) a vizsgálóknál azonban igen költségesek és igen kevésbé hatékonyak, míg a kapcsolók (links) nem költségesek, viszont hatékonyak. Ez azért van, mert a szerepjelölő önmagában monadikus, szemben az egybekapcsolt szubsztantívumok diadikus természetével. A szerepjelölők és a kapcsolók együttes alkalmazásá-

nál flexibilisebb artikulációt eredményeznek az un. szupraszintaktikus elemek (amelyek hasonlítanak a kapcsolókhöz).

Minden elemző, jelzetelő és tárolás-stratégiai elméletnek és gyakorlatnak egy közös célja van: az információ hatékony visszakeresése, mégpedig annyira leegyszerűsítve, hogy a számítógép legyen képes elvégezni a visszakeresési aktust. Az erre szolgáló elemzés mellesleg döntően befolyásolhatja a kézi visszakeresési apparátusokat is. A visszakeresés teljes mértékben függ a tárolás stratégiájától és ezt megelőzően a tárolás és visszakeresés fogalmi szervezésétől. Ennek olyannak kell lennie, hogy megtaláljuk a relevánsat, azaz azt, ami segítségünkre szolgál. Grolier, Gardin, Hjelslev és a saját (kategóriák és relátorok) rendszerének alapjait egybevetve, arra a különös tapasztalatra jut a szerző, hogy bármennyire különbözőznek is a nézőpontok egymástól, abban megegyeznek, hogy a tripartíció (hármasszós) közös bennük. (Mellesleg az is érdekes, hogy az összehasonlítás lehetséges a természetes nyelvi és a mesterséges nyelvi osztályozási rendszerek között.)

Ism. HORVÁTH Magda