

AZ INFORMÁCIÓ-VISSZAKERESÉS PROBLÉMÁI:
MŰVELETEK A DOKUMENTUMOK TÁRGYÁT FELTÁRÓ FOGALMAKKAL

Hoványi Gábor

A bibliográfia, az irodalomkutatás és az információ-visszakeresés feladatai

A szakirodalmi tájékoztatásnak működése szempontjából két fő formája van:

1. a még nem dokumentált, elsődleges ismeretanyag /könyvek, folyóiratcikkek stb./ folyamatos dokumentációs feltárása és közreadása;
2. esetenként megadott kérdésekre szakirodalom összeállítása.

Ennek az összeállításnak anyaga elsősorban a már dokumentált, tehát a másodlagos ismeretanyag; ezt általában csak kiegészíti az az irodalom, amelyet a feltáráskor nem dolgoztak fel - vagyis amelyre csak a megadott kérdések szakirodalmának összeállításakor irányul a kutató figyelme. /Ilyenek pl. az ún. "rejtett bibliográfiák"./ Az összeállítás alapvető formája pedig a szakirodalmi jegyzék, amelyet bibliográfiának vagy irodalomkutatásnak neveznek.

A bibliográfia, mint köztudomásu, eredetileg könyveket leíró, azaz felsoroló jegyzéket jelölt. Az újabb típusu kiadványok /folyóiratok stb./ megjelenése után azonban bibliográfiának neveztek azokat a jegyzékeket is, amelyek ezeknek a kiadványoknak a közleményeit irták le a könyvtári módszerekkel. Így alakult ki végül a bibliográfiának olyan fejlett formája, mint pl. az annotált szakbibliográfia. /Az időszaki bibliográfia viszont a szakirodalmi tájékoztatás első formájával, azon belül a referálólappal rokon./

A szakirodalmi tájékoztatás felhasználja a bibliográfiák készítésének szinte valamennyi elméleti és gyakorlati eredményét, de arra törekszik, hogy elhatárolja magát a bibliográfiától mint műfajtól: szakirodalmi jegyzékeit is inkább "irodalomkutatásoknak" nevezi. Az idegenkedés fő oka aligha lehet egy olyan formai kérdés, mint az, hogy az irodalomkutatás a bibliográfiával ellentétben nem az eredeti, hanem /csak vagy első helyen/ az olvasó nyelvén sorolja fel szakirodalmi címeit. A megkülönböztetést inkább tartalmi szempontok miatt kívánják: a kifejezetten szakmai jellegű irodalomkutatások összeállítása az anyag alapos tartalmi ellenőrzését követeli meg, ezt pedig

általában nagyobb hozzáértéssel végzi az adott szakma szakembere, mint a szakmában csak többé-kevésbé otthonos könyvtáros. A feltárás mélységének oldaláról ez azt jelentené, hogy a bibliográfiának inkább indikatív /jelző/, az irodalomkutatásnak inkább informatív /tájékoztató/ a szerepe. A feltárás módja szerint pedig úgy lehetne jellemezni a kettőt, hogy a bibliográfia főként az elsődleges ismeretanyag tartalmát kifejező egy vagy több fogalmat csak egységes egészként tartja számon /pl. az osztályozás 2-3 viszonyított jelzetével/, az irodalomkutatás viszont a másodlagos ismeretanyag alapvető fogalmaira is épül /pl. a referálólapok indexeire, amelyek referátumokként 4-8 tárgyszót is tartalmaznak/. Ezek a jellemzők azonban nem mindig és nem maradéktalanul érvényesek, és talán ezért sem tudtak eddig megállapodni a könyvtári és a tájékoztatási szakemberek a két műfaj éles határvonalában.

A bibliográfia és irodalomkutatás elnevezésekhez az utóbbi években egy harmadik rokon-fogalom kapcsolódott: az információ-visszakeresés. Ezt viszont három sajátossága élesebben elhatárolja a bibliográfiától és az irodalomkutatásoktól.

A szakemberek - különösen pedig a műszaki szakemberek, a kutató-, tervező- és szerkesztőmérnökök - egyre aprólékosabb kérdésekre várnak választ a tájékoztatástól. A válaszok ezért egyre ritkábban tartalmazhatnak olyan elsődleges ismeretanyagot, amelynek egésze a feltett kérdésre vonatkozik /bibliográfiai /?/ feltárási mód/. De mind többször eredménytelen az a kutatás is, amely a másodlagos ismeretanyag egységeként 4-8 alapvető fogalma között keresi a választ /irodalomkutatási /?/ feltárási mód/. Ezekkel szemben az információ-visszakeresés általában az elsődleges ismeretanyag egységeként 10-50 alapvető fogalmát használja fel a válaszadásban. Ez az információ-visszakeresés első megkülönböztető sajátossága.

A szakembereket azonban nemcsak az irodalomban rögzített ismeretanyag érdekelheti: pl. hangszalagon rögzített előadások vagy új termékeket bemutató műszaki filmek részletei is hasznosak lehetnek számukra. Ebből a problémából született az a felismerés, hogy a nem irodalmilag rögzített ismeretanyag tájékoztatási feltárását, tárolását és visszakeresését is az információ-visszakeresés alapvető elvei szerint lehet végezni. Ez az információ-visszakeresés második megkülönböztető sajátossága.

Harmadik megkülönböztető sajátosságának megemlítése előtt azonban - mintegy érvként is - tekintsük át működése egyik szakaszának főbb fázisait.

Műveletek a dokumentumok tartalmát feltáró fogalmakkal

Az információ-visszakeresésnek a dokumentumokat két szempont szerint kell leírnia: rögzítenie kell azok bibliográfiai adatait, hogy a dokumentumokat meg lehessen találni, és fel kell tárnia a tartalmukat, hogy egy megadott tárgykör-

re a megfelelő dokumentumokat lehessen megtalálni. Az első fajta leírást általában a szokásos dokumentációs eljárással végzi; a tartalom feltárása pedig végső soron fogalmakkal történik. Az információ-visszakeresés egyik alapvető elméleti feladata annak a meghatározása, hogy a dokumentumok tárgyát feltáró fogalmakkal milyen műveleteket, milyen mértékben kell végeznie ahhoz, hogy a visszakereséskor elérje a kívánt hatásfokot. A következőkben a fogalmakkal végezhető fontosabb művelet-típusokat mutatjuk be.

Helyettesítés

Egy ismeretanyag tartalmát feltáró fogalmak leírása az ismeretanyag szövegéből választott jellemző szavak kiemelésével vagy a szövegnek csak értelmétől, de nem a formájától függő szavakkal történhet. Ezt a két fogalomleíró szótípust tárgyszónak nevezzük. A csak az ismeretanyag formájától vagy értelmétől függő tárgyszavaknak igen nagy hátránya, hogy formailag is és értelmileg is labilisak: példák erre a szinonimák /autó-gépkocsi/ és a többjelentésű szavak /pl. a "rác" ásványtani, fénytani, kohászati, építészeti és híradástechnikai jelentése/ lehetnek. Labilitásuk miatt az információ-visszakeresés csak clyan ritka esetekben használja a tárgyszavakat a tartalom feltárására, amikor ezt a hátrányt más szempontból jelentős előny feledteti: a szövegkörnyezetes címtárgyszavas mutató /KWIC/ esetében pl. a gépi feldolgozás gyorsasága. Ennél a mutató-típusnál a címtárgyszó, vagyis az a tárgyszó, amelyik megszabja a mutatóba-sorolás helyét, eredeti szövegkörnyezetével együtt szerepel a mutatóban; a címtárgyszó első jelhelyétől jobbra és balra szerepeltethető jelek száma kötött /1./

/1. ábra/. Kérdés azonban, hogy a címtárgyszavas mutatónak lesz-e jelentősebb szerepe az információ-visszakeresésben, hiszen elsősorban címek, esetleg referátumok indexelésére alkalmas; általában csak ezek szövegében következnek olyan szoros egymásutánban a tárgyszavak, hogy a kötött jelszámba beleesve érdemben kiegészíthetik a címtárgyszó értelmét.

1. ábra

a tárgyszavak labilitásának káros hatását. Ezért ezeket meghatározott és tervszerűen fejlesztett fogalom-gyűjtemény szavaival, az ún. vezérszavakkal helyettesítik. Ezt mutatja a 2. ábra.

A helyettesítéssel el lehet érni, hogy egy fogalom a feltáráskor és a visszakereséskor mindig ugyanazzal a szóval és annak ugyanabban a formájában szerepeljen. A vezérszójegyzék összeállítását vagy összefoglaló jellegű szakkönyvek fogalomanyagának /tárgymutatójának/ összedolgozásával, vagy meghatározott számú dokumentum indexe-

THE USE OF ULTRASONIC IN MELTS AT HIGH TEMPERATURES.
APPLICATION OF ULTRASONIC IN NUCLEAR MEMBRANE AS A BARRIER TO THE FLOW OF
MEMBRANE - SELECTIVE MEMBRANE CELLS FOR THE PREPARATION O
CATION IN MITOCHONDRIAL MEMBRANE FRAGMENTS.
ACIDS IN THE VITELLINE MEMBRANE OF A LARVA OVUM.
SELECTIVE BIOLOGICAL MEMBRANE TRANSPORT OF INDIVIDUAL AMI
THE PLACENTA AND FETAL MEMBRANES OF THE RABBIT.
PERMEABILITY OF CAPILLARY MEMBRANES.
PERMEABILITY OF LFTOP ASHIC MEMBRANES.
PERMEABILITY OF THE CELL MEMBRANES BY ELECTRO-OSMOSIS.
MEMBRANES OF TUNA OILS.
MEMBRANES.
MEMBRANES IN NORMAL
CELLS OF THE SOLVENT IN
CELLULAR MATERIALS. 2- MENTHADENOL OF THE ESSENTIAL OILS O
RELATIVES AND ANALOGUES OF MENTHADENOL.
IN VITRO METABOLISM OF MENTHADENOL 2-0- TOLYLUY-112- PROP
MERCAPTOPYLAMMONIUM SALTS AND THEI

HELVOL-80-177
HESSVU-80-485
LQJW-80-988
KARFLM-80-MPH
PAKEL-80-LSC
SMALV-80-SFF
BRASDM-80-SBN
TERKI-80-IEA
GAGDM-80-ATP
MUTHE-80-IES
JANVU-80-TRE
PEIFJ-80-DLN
LJRECH-80-AGE
GONJMG-80-RSH
BAVETR-80-JPH
PETROV-80-OBP
HAASAR O-VNH
BOMHEN-80-VMS

Tárgyszavak	Vezérszavak
autó	gépkocsi
lámpák	fényszóró
beállítás	beállítás
készülék	műszer
...	...

2.ábra

A tárgyszavak helyettesítése vezérszavakkal

lésével végzik. A nyers vezérszavakat ezután több menetben vetik össze, hogy a többértelműséget és a fogalomátfedéseket kiszűrjék. Szélesebb szakterületeken a vezérszavakról betűrendes és osztályozott jegyzéket is készítenek; a betűrendes jegyzékbe a vezérszóként el nem fogadott tárgyszavakat is be lehet építeni, de ekkor ezekről utalni kell a megfelelő vezérszóra. Ez az egyik elvi alapja a nem tárgy-, hanem vezérszavas gépi feltárásnak is.

Kapcsolás

A 2. ábrán bemutatott egyes vezérszavak azonban kevés érdemlegeset mondanának a visszakereséskor, mert önmagukban túl tág a jelentéstartományuk. Egyre inkább csak a több vezérszó összekapcsolásával szűkített jelentéstartomány felel meg a visszakeresés igényeinek. A feltárást és a visszakeresést ezért általában több szóból álló vezérfelfűzéssel végzik.

Az információ-visszakeresésben azonban alapos elméleti és gyakorlati munkával olyan rendszert is kifejlesztettek, amelyik nem szorul a nehezkesebb vezérfűzésekre, hanem csak egytagú vezérszavakkal dolgozik. Ez az Egyvezérszavas /Uniterm/ rendszer. /2/ Lényege, hogy nem dokumentumonként, hanem vezérszavanként tárolja a feltárt és egy előzetes nyilvántartásban dokumentum-számmal jelzett ismeretanyagot /3. ábra/.

3.ábra

Műszer									
Beállítás									
Fényszóró									
Gépkocsi									
1	2	3	4	5	6	7	8	9	0
6711				4525	1366			7939	
8931				5135					
				5345					

Az Egyvezérszavas /Uniterm/ rendszer kézi feldolgozás esetén

/A kézi visszakeresés megkönnyítésére a dokumentumszámok utolsó számjegyük értéke szerint 10 oszlopba sorolva szerepelnek/

A visszakeresés a választott vezérszavakhoz bejegyzett dokumentumszámok egyeztetésével történik.

Az Egyvezérszavas rendszernek több gyenge pontja van: példaként elég bemutatni, hogy a "Gépkocsi fényszórók tervezése" témára a "Gépkocsi fényszóró-beállító műszereinek tervezése" tárgyú dokumentumot is ki fogja irni, mert a keresett első tárgykör valamennyi vezérszavánál szerepel a második tárgykör dokumentumszáma.

Ez a fogyatékoság is azt kívánja, hogy az egy fogalomkörű alkotó vezérszavak egy vezérkifejezésben szerepeljenek: erre szolgál a kapcsolás művelete. A három alapvető kapcsolás a szabad-kapcsolás, a rendező-kapcsolás és a kötő-kapcsolás.

A szabad-kapcsolásban a vezérszavak kapcsolódása és sorrendje a vezérkifejezésben az értelemhűségein belül bármilyen lehet. /A magyar nyelvben a szórend nem igen okoz olyan értelemváltozást, mint a klaszikus angol példában: the blind Venetian = a vak velencei; the Venetian blind = az ablakközi lécroletta./ A szabad kapcsolásra példa a szokásos betürendes mutató elve többszörös besorolás esetén /4. ábra/.

Tárgykör: Kerámiai anyagok felületi ellenállása		
Betürendes mutató vezérkifejezései:	1	Anyag/ok/, kerámiai, felületi ellenállás/a/ ...
	2	Ellenállás, felületi, kerámiai anyag/oknál/ ...
	3	Felületi ellenállás kerámiai anyag/oknál/ ...
	4	Kerámiai anyag/ok/ felületi ellenállás/a/ ...

4. ábra

Változatlan értelmű szabad kapcsolás a betürendes tóban, többszörös besorolás esetén

A "Salétromsav hatása hexaminra RDX /nagyhatású robbanóanyag/ előállításakor" tárgyi dokumentum vezérkifejezésében szabad-kapcsolás esetén az RDX - hexamin - salétromsav vezérszavak bármilyen sorrendben követhetnék egymást. Ennek viszont az lenne a következménye, hogy a visszakeresés nem tudna különbséget tenni a

hexamin + salétromsav = RDX, valamint az
RDX + salétromsav = hexamin között, és

mindkét tárgykörre kiírná az említett dokumentumot. Ezt a zavart szűri ki a rendezőkapcsolás azzal, hogy szabályokhoz köti a vezérszavak kapcsolásának sorrendjét. A legismertebb rendező-kapcsolás az alkotó-elemzésen /facet-analysis/ alapuló osztályozás elve: ez fogalomkategorizációt állít fel /pl. alanyi szubsztanciák, ezen belül természetiek és termékek, műveletek, művelet-tárgyak stb./, az egyes ka-

tegóriákba sorolható vezérszavakat azután a kategóriák megszabott sorrendje szerint kell kapcsolni. A kettőspontos Osztályozás eredetileg öt rangnathani kategóriáját /3/ az információ-visszakeresés céljára VICKERY 25 tagu kategória-rendszerre fejlesztette tovább./4/ Egy a hajtóművek gyártásával foglalkozó közleményben azonban szó eshet például a fogaskerekek lángezéséről és a forgattyustengelyek méretezéséről. Az ezt feltáró vezérszavak a következők lesznek:

1. fogaskerék
2. lángezés
3. forgattyustengely
4. sörétfuvás

Az 1. és 3., valamint a 2. és 4. vezérszó ugyanabba a fogalomkategóriába esik. Ha csak a kategóriák között különbséget tévő rendezőkapcsolással tárjuk fel a dokumentumot, a visszakeresés a fogaskerekek sörétfuvása vagy a forgattyustengelyek lángezése tárgykörre is kiirhatja a közleményt. Ezt akadályozza meg a kötőkapcsolás, amikor megszabja, hogy egy dokumentum vezérszavai közül melyek kapcsolódhatnak egymással. Ennek példája az Egyvezérszavas rendszer egyik továbbfejlesztése, /5/ amelyben a dokumentumszámok után, ponttal elválasztott számmal jelzik azokat a fogalmi szinteket, amelyeken a kapcsolásokat végezni szabad. Ezt mutatja az 5. ábra.

	Szabad kapcsolás		Kötőkapcsolás	
vezérszavak	fogaskerekek	lángezés	fogaskerekek	lángezés
dok. szám	64955	64955	64955.1	64955.1
vezérszavak	forgattyusteng.	sörétfuvás	forgattyusteng.	sörétfuvás
dok. szám	64955	64955	64955.2	64955.2

5. ábra

A kötőkapcsolás elve és megvalósítása az Egyvezérszavas rendszerben

Szerepjelölés

BAR HILLEL a szakirodalomban sokat idézett példája a "Németország borexportja Franciaországba" /6/: ennél sem a kategóriákon alapuló rendezőkapcsolással /Francia- és Németország ugyanaz a kategória!/, sem a kötőkapcsolással /Franciaország és Németország ugyanaz a fogalmi szint!/, nem lehet meghatározni, hogy a borexport Franciaországból Németországba, vagy Németországból Franciaországba irányul-e? Az ilyen típusú feladatokat az információ-visszakeresésben a szerepjelölés műveletével oldják meg. Az egyik legrészletesebben kidolgozott visszakereső rendszer,

a Western Reserve University /WRU/ rendszere, huszonöt narombetűs kóddal jelzi a vezérszavak fogalmainak szerepét. /A szerepjelölő kód megkülönböztetésére az első betű mindig K./ A dokumentumok feltárássakor kétszlopos nyomtatványt töltenek ki: a jobboldali oszlopba a tárgyszavakat írják be, majd a baloldaliba a tárgyszavaknak a tárgy kifejezésben betöltött szerepét meghatározó szerepjelölőket vezetik be. Egy ilyen feltárás egyszerűsített formáját mutatja be a 6. ábra /7/; a szerepjelölők kódja mellett zárójelben - a példa érthetővé tételére - azok jelentését is feltüntettük.

KEJ	/művelet alapanyaga/	7075-0 Al ötvözet
KAM	/művelet/	robbantásos alakítás
KQJ	/művelet eszköze/	odór
KUJ	/főkomponens/	4130 acél
KAH	/művelet feltétele/	nyomás
KAH	/művelet feltétele/	magas hőmérséklet /880°F/
KWJ	/művelet eredménye/	ürhajtó-ajtó
KUP	/meghatározott tulajdonság/	huzószilárdság
KUP	/ " /	folyási határ
KUP	/ " /	nyulás

6.ábra

Szerepjelölés a WRU visszakereső rendszerében

A szerepjelölésnek azonban ennél az elektronikus adatfeldolgozó gépre méretezett eljárásnál egyszerűbb megoldását is ki lehet alakítani. Az Egyvezérszavas rendszer említett továbbfejlesztésében /5/ pl. a dokumentumszám mellé irt A...J betűvel határozzák meg a vezérszó szerepét.

Elemzés

A kapcsolat és a szerepjelölés művelete lehetővé teszi az egy fogalmat leíró vezérszavakból az összetett fogalmat leíró vezérfelvezetések kialakítását. Az információ-visszakeresésben azonban vannak olyan látszólag egyetlen fogalmat leíró tárgyszavak, amelyeket célszerű elemi fogalom-alkotókra bontani. Ez a feladat az elemzés műveletével végezhető el.

A "hőmérsékletmérő" tárgyszó pl. látszólag egyetlen fogalmat jelöl. Voltaképpen azonban három elemi fogalom-alkotóra lehet bontani: a hőmérsékletmérő 1. eszköz, amelyet 2. hőre vonatkozó 3. mérésre használnak. Az elemzés ebben az esetben visszakereshetővé teszi az eredetileg "hőmérsékletmérő" tárgyszóval jellemzett anyagot pl. a "mérőeszközök" vagy a "hőmérők" témára is.

Ezt az elvet fejlesztette tovább a WRU szemantikailag elemző rendszere. /8/ Ebben a "hőmérsékletmérő" tárgyszó három fogalom-alkotójának megfelelő vezérszavak kódolása a következő:

eszköz: M.CH
 mérés: M.SR
 hő: R.HT, ebből hőmérséklet R.HT 1004.

A kódok hiányzó második jelhelyére a rendszer a szerepjelölés műveletével illeszti az A,U,W kódot, amelyek arról tájékoztatnak, hogy a hőmérsékletmérő az eszköz /MACH/, amelyet hőmérsékletnek a /RWHT 1004/ mérésére /MUSR/ használnak. A hőmérsékletmérőkön belül az egyes típusokat úgy különböztetik meg, hogy még egy négyjegyű számot kapcsolnak a kódhoz. A szerepjelöléssel kiegészített elemzés végeredményének szemléltetésére és az elemzés szerepének érzékeltetésére bemutatjuk a következő, látszólag egy-egy nem összetett rokonfogalmat leíró tárgyszavak WRU-Kódját:

Hőmérő /kaloriméter/	MACH	MUSR	RWHT	3001	
Hőmérsékletmérő /termométer/	MACH	MUSR	RWHT	1004	3002
Hőmérsékletjegyző /termográf/	MACH	MUSR	RWHT	1004	3007

Az elemzés műveletének a hozzáférések számának növelésén kívül még egy előnye van: a különösen a műszaki szakterületeken viszonylag számos fogalmat kevés elemi alkotóval írja le. A WRU rendszerben pl. a kohászat anyagának 38 000 fogalmát az elemzés segítségével 237 elemi fogalom-alkotóval tudták kifejezni. /7/ Az elemzés különleges, de egyszerű esete az értelmező elemzés. Ez a vezérszóhoz az értelmét szűkítő fogalmat fűz, pl.

rács, híradástechnikai
 rács, ásványtani stb. formában.

Rendelés

Az elemzés művelete a dokumentumot leíró fogalmakat alkotóikra bontja. Fordított művelete a rendelés, amely a fogalmakat meghatározott szempont szerint növekvő vagy szűkülő jelentéstartományuk viszonyításával főlé-, ill. alárendeli egymásnak. A rendelés tehát megegyezik a könyvtári osztályozás legalapvetőbb feladatával; az információ-visszakeresésben mégis helyesebb a rendelés szó használata, mert ez a könyvtári osztályozás sokrétűbb feladatköréből csak a főlé- és alárendelést jelöli.

A rendelésnek három főfeladata van az információ-visszakeresésben:

1. a pusztán formai besorolásokkal /pl. betűrendes mutató/ szemben tartalmi vezérfonal a feltáráshoz és/vagy a visszakereséshez szükséges vezérszavak megtalálására;
2. lehetővé teszi, hogy az alárendelt fogalmakkal feltárt dokumentumot a főlérendelt fogalom iránt érdeklődőknek vissza lehessen keresni;
3. módot ad arra, hogy egy vezérszóval végzett eredménytelen visszakeresést a közvetlenül főlérendelt vezérszóval lehessen megismételni - ha ez is segítséget nyújt az érdeklődőknek.

	JÁRMŰIPAR		VEGYIPAR	
Vezérszó és ETO-jelzetek	gumiköpeny	629.11.012.553.2	gumiköpeny	678.76.066
1. főlérendelt fogalom	köpeny	629.11.012.553	műgumik felhasználása	678.76.06
2. " "	légabroncs	629.11.012.55	műgumik	678.76
3. " "	abroncs	629.11.012.5	szintetikus uton elő- állított polimerizá- ciós termékek	678.7
4. " "	futómű részei	629.11.012	gumi és műanyagipar	678
5. " "	járműrészek a ha- jóművek kivételé- vel		különböző iparágak	67
6. " "	közuti járművek	629.11.01	alkalmazott tudományok	6
7. " "	közlekedéstechnika	629.11		
8. " "	műszaki tudományok	629.1		
9. " "	alkalmazott tudományok	6		

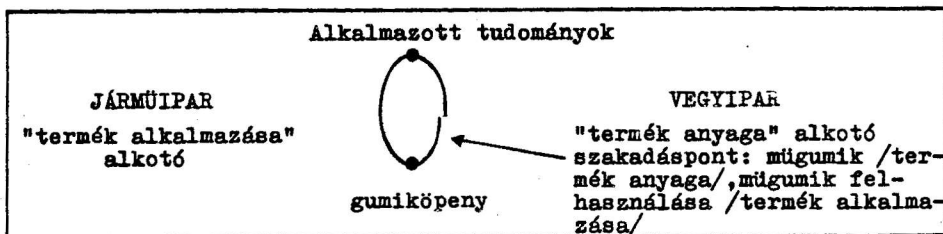
7. ábra

Azonos vezérszó szakterületenként különböző főlérendelt fogalmak

A rendelés önmagában nem egyértelmű művelet: azonos vezérszó fölé-, ill. alárendelt fogalmi szakterületenként különbözhetnek. Példaként a "gumiköpeny" /gépkocsi/ vezérszó ETO szerint fölérendelt fogalmait a jármű- és a vegyiparban a 7. ábra mutatja.

Miként a példa elemzéséből is kitűnik, az azonos vezérszónak fölérendelt egyes fogalmak az alkotó-elemzés /facet analysis/ kategóriái szerint válnak szét: a járműiparban a "termék alkalmazása" alkotó szerint történik a fölérendelés, a vegyiparban pedig többnyire /lásd később/ a "termék anyag" alkotó szerint. Általánosítható, hogy:

1. a rendelést csak a szempontját meghatározó alkotó teszi egyértelmű műveletté;
2. minden többalkotós fogalom azonos vezérszóként fogható fel;
3. azonos vezérszó esetén az alkotók szerint végzett fölérendelések olyan folyamatos görbéket írnak le, amelyek végül az egyik fölérendelésnél metszik egymást, azaz egy fogalomban egyesülnek /8. ábra/;
4. azonos vezérszavak nem az alkotóik szerint végzett fölé- és alárendelések nem kapcsolódnak folyamatosan, vagyis görbéjük a metszéspontjuk előtt megszakad /lásd a 8. ábrát/.



8. ábra

Azonos vezérszó alkotók szerint különböző fölérendelése

A megszakadó görbe a gyakorlatban azt jelenti, hogy a rendelés művelete három főfeladatának csak részben vagy egyáltalán nem tud elegendet tenni. Ezért az információ-visszakeresés rendelési műveletében - ill. az ahhoz szükséges osztályozórendszer kialakításakor - hasznos a Kettőspontos osztályozás következetes alkotóelemzési módszerének /9/ az alkalmazása.

Magát a rendelés műveletét vagy a feltáró- és visszakereső rendszer egészében, vagy annak egyes részeiben lehet végezni. Az utóbbira, vagyis a részleges rendelésre példa az olyan betűrendes vezérszórendszer, amelyben egyes vezérszavakhoz a betűrenden belül újra kezdődő betűrendben besorolják az adott vezérszavak alárendelt fogalmait /9. ábra/.

Paramágneses rezonancia /lásd Mágneses magrezonancia/
 Parázsfénykibocsátás
 - benzolgőzben 1627
 -,Disszociáció hatása -re 2095
 - molekuláris gázokban 1442
 - spektrométeres vizsgálata 1780
 Paritás 127

9.ábra

Betűrendes mutató részleges rendeléssel
 /"bokrosítás"/

Sulyozás

A sulyozás célja annak a meghatározása, hogy a dokumentum egyes vezérszavai milyen mértékben tárják fel a dokumentum tartalmát. Az információ-vizsszakeresésben több módszert is javasoltak a vezérszavak sulyozására. Az egyik módszer /10/ szerint például a feltáráskor 1/8-8/8 értékű, meghatározott követelménnyel jellemzett csoportba kell sorolni a dokumentum vezérszavait. A vezérszavak sulyértékeinek összevetése képet ad a dokumentum vezérszavakkal jelölt tartalmi sulypontjairól is. Például a "Légkör /3/8/, Élelem /6/8/, Ember /2/6/, Anyag /3/8/, Ürrepülés /6/8/ és Csomagolás /7/8/", vezérszavak közül a Csomagolás- Élelem- Ürrepülés vezérszavak jelölnék a dokumentum tartalmi sulypontját.

A sulyozás különösen gépi visszakeresés esetén könnyítené meg az eligazodást a visszakeresett anyagban. Ekkor ugyanis programozható lenne, hogy kiírásakor milyen vezérszavak milyen sulyértékű és milyen kapcsolású változataival leírt dokumentumok milyen sorrendben kövessék egymást. Ezzel nemcsak a visszakeresést, hanem a visszakeresés "értékelését" is gépesíteni lehetne.

A sulyozást - jogosan vagy kevésbé jogosan - lehet pusztán mennyiségi feladatnak is tekinteni. Ekkor a vezérszavaknak megfelelő tárgyszavak /visszahelyettesítés/ sulyát előfordulásuk száma és mennyiségileg jellemezhető helyzete szabja meg. Ezzel a sulyozás gépesíthető műveletté válik, hiszen problematikájának alapja azonos a gépi referátum készítésével. Erre pedig már több elméletet is kialakítottak, /11/ és a gyakorlatával is kísérleteznek. Tény azonban, hogy valamennyi tárművelete közül máig a sulyozás kristályosodott ki a legkevésbé.

Besorolás

A dokumentumok tartalmát feltáró fogalmakkal utoljára a besorolás műveletét kell elvégezni: a vezérszavakat és az ezekhez kapcsolódó adatokat meghatározott rendben be kell sorolni a tárba, hogy onnét kívánságra visszakereshetők legyenek. A táarak sok esetben meghatározzák a besorolás mikéntjét: más követelményeket

állít fel egy könyvformában megjelenő betűrendes tárgymutató, és másokat egy elektronikus számológép mágnesszalagos memóriája. A besorolás műveletének néhány alapvető formája azonban a tár-típusoktól függetlenül is rögzíthető.

C é l j e szerint a besorolás fogalom-leíró és dokumentum-leíró lehet. Az első a dokumentum tartalmát feltáró vezérszavakat vagy vezérkifejezéseket külön-külön sorolja be a tárba; a második a dokumentumot feltáró valamennyi vezérszót vagy vezérkifejezést együtt sorolja be. Fogalomleíró besorolás pl. az Egyvezérszavas rendszeré; dokumentum-leíró pedig a Keresőoszlopos mutatóé /10. ábra/. Ennél az egyes dokumentumszámokhoz sorolják be a dokumentum valamennyi vezérszavának jelzetét: a kétszámjegyes jelzetcsoportok megfelelő oszlopába bejelölik az adott jelzetcsoportba tartozó egyes vezérszavak jelét.

Dok.számok	11	12	13	14	15	16	17
01101		N		m/%		?Z	
01102	6			je	mv		
01103	Y		gm				
01104		7		kwj		;5	

10. ábra

Keresőoszlopos mutató

A fogalomleíró besorolás előnye, hogy annyi hozzáférést tesz lehetővé a visszakeresés számára, ahány fogalom tárja fel a dokumentumot. Legfőbb hátránya viszont, hogy még kötőkapcsolás esetében is csak arra tud válaszolni: dokumentumként milyen visszakeresett fogalmak kapcsolódnak egymáshoz - de adós marad azzal, hogy milyen tárolt, de vissza-nem-keresett fogalmak kapcsolódnak még a visszakeresett fogalmakhoz. Ezért a dokumentumról a maga egészében csak a dokumentum-leíró besorolás adhat képet. Ennek hátránya viszont az, hogy egyszeres besorolás esetén visszakereséskor az egész tárat végig kell keresni; többszörös besorolás esetén pedig nagy a besorolás idő- és a tár elvesztése. /A legújabb gépi megoldások ezt a hátrányt jelentősen csökkentik azzal, hogy egyszerre több leolvasófejvel keresnek vissza./

T á r g y a szerint a besorolás vezérszóbesorolás vagy vezérkifejezés-besorolás lehet: az elsőre az Egyvezérszavas rendszer, a másodikra a betűrendes mutató részleges rendelése a példa.

R e n d j e szerint a besorolás általában betű-, szak- vagy számsorrendes lehet. Szélesebb szakterületek nagyvolumenű anyagának betűrendes besorolásához szakrendes kiegészítő mutatót, szakrendi besoroláshoz pedig betűrendes kiegészítő mutatót készítenek. Ezek jelentőségére hívja fel a figyelmet a legelterjedtebb visszakereső rendszerek cranfieldi összehasonlító értékelése. /12/ Ez kidomborít-

ja, hogy egy olyan "kitaposott utakat járó" szakrendszer, mint az ETO, az értékelés során nem kis mértékben betürendes mutatójának köszönhető versenyképességét.

M ó d j a szerint a besorolás egyszeres, többszörös vagy utaló lehet. Az egyszeres besorolás egy vezérszót vagy vezérkifejezést csak egyszer sorol be. Ilyen pl. a WRU visszakereső rendszerének a besorolása: ebben az egy dokumentumot feltáró valamennyi vezérkifejezést, ill. annak "mondatokká" formált különleges egységét a tár egy helyére sorolják be. A tárból GE 225-ös típusu elektronikus számológéppel egy menetben 8 óra alatt 20 témára keresik vissza a tárolt 100 000 dokumentumból az anyagot. A párhuzamos visszakereséssel is ellensúlyozni tudják azt a hátrányt, hogy visszakereséskor az egész tárat végig kell olvasni. /13/ A többszörös besorolás végső célja, hogy dokumentumonként annyi hozzáféréssel tegye lehetővé a visszakeresést, ahány vezérszóval, ill. vezérkifejezéssel a dokumentum feltárása történt. Többszörös vezérszó-besorolás az Egyvezérszavas rendszer; többszörös vezérkifejezés-besorolás "mechanikus" megoldása pedig a körforgó besorolás /ll. ábra/. Ebben a vezérkifejezések vezérszavai folyamatosan változó sorrendben követik egymást; mindig a sorrend-váltáskor élrekerülő vezérszó a címvezérszó, azaz ez szabja meg a besorolás helyét a tárbán.

A besorolás példája		elve
<u>Szuperszonikus</u>	áramlás nyomása szárnyfelületre, 1439	ABCD
<u>Áramlás</u>	nyomása szárnyfelületre, szuperszonikus, 1439	BCDA
<u>Nyomás</u>	szárnyfelületre, szuperszonikus áramlásé, 1439	CDAB
<u>Szárnyfelületre</u>	szuperszonikus áramlás nyomása, 1439	DABC

ll. ábra

A körfogós besorolás

/Az aláhuzottak a címvezérszavak, a vezérkifejezések után a dokumentumszám következik/

A körfogós besorolás azonban nem gazdaságos művelet, ha rendezést is kell végezni. Ekkor ugyanis a vezérkifejezés "n" számú vezérszavából $n / 1 + n /$ besorolást kellene végezni. Ha $n = 3$, akkor gazdaságosabb az ² alkotó-elemzés láncolatos besorolása, amelyiknél a besorolások száma $2n$. /14/ A láncolatos besoroláskor a címszóként már szerepelt vezérszavakat a sorrend-váltáskor törölni kell a vezérkifejezésből, az egyes vezérkifejezések után pedig a vezérszavak fordított sorrendjében rendelésük jelzetét kell feltüntetni. A besorolás az egyes címvezérszavak és a vezérkifejezésüknek megfelelő rendelés-jelzetek szerint történik /12. ábra/.

A besorolás műveletének különleges formája az utalás. Az egyenes utalót, amely a vezérszó rokonfogalmáról utal a vezérszóra, címtárgyszóként kell besorolni. A bővítő utalót, amely egy értelmileg bővíthető vezérszóról egy vagy több bővítő vezérszóra utal, a bővítő

hető vezérszóhoz kell besorolni. Keresztutaláskor a bővítő utalót fordított utalással a bővítő vezérszavakhoz is be kell sorolni.

Vezérkifejezések	Rendelés elve
Szuperszonikus /A/, áramlás /B/, nyomás /C/ szárnyfelület /D/ Áramlás /B/, nyomás /C/, szárnyfelület /D/ Nyomás /C/, szárnyfelület /D/ Szárnyfelület /D/	DCBA DCB DC D

12. ábra

Láncolatos besorolás
/A vezérszavak után zárójelben a rendelés elvi jelzete szerepel/

A jelentősebb visszakereső rendszerekkel végzett műveletek. Végkövetkeztetések

A dokumentumok tárgyát feltáró fogalmakkal a felsorolt hét alapvető művelet végezhető el. Ezeket átfogóan különböző információ-visszakereső rendszerek példáin ismertettük. Az egyes rend-

szerekkel azonban nem végeznek el minden műveletet: vannak rendszerek, amelyekkel egyes műveletek vagy műveletformák nem is végezhetők el: ilyen pl. a dokumentumleíró besorolás az Egyvezérszavas rendszer esetében. Az 1. táblázat mutatja, hogy melyik visszakereső rendszerrel általában milyen műveleteket végeznek.

Mindezek után felmerülhet a kérdés: a jelentősebb visszakereső rendszerek közül melyik a leghatékonyabb? Erre nem lehet egyértelmű választ adni: a hatékonyságnak olyan külső tényezői vannak, mint a feldolgozott anyag volumene, a feltárás megkívánt mélysége, a visszakeresés gépesítésének foka stb.; a belső tényezők közül pedig csak egyik a tartalmat feltáró fogalmakkal végezhető műveletek szerepe. A hatékonyság kérdése ezért önálló problémakör, amelyet egészében helyes vizsgálni. Ennek a vizsgálatnak pedig csak egyik részfeladata lehet annak rögzítése, hogy milyen dokumentációs anyaggal milyen típusú feladatokra, melyik műveletet vagy műveletsort célszerű és gazdaságos egy információ-visszakereső rendszerben elvégezni.

A két alapvető művelet ismertetéséből azonban két következtetést is le lehet vonni.

A bibliográfia-készítéshez főként formai szabályokat állítottak fel; az irodalomkutatások összeállításához ezeken túlmenően főként azokat a segédeszközöket és felhasználásuk módszerét jelölték meg, amelyeket a szakirodalmi tájékoztatás szolgáltatásaival hozott létre. Az információ-visszakeresés viszont kezdi összefüggésükben

Műveletek	Visszakereső rendszerek										
	Alkotóelemzés	Betűrendes mutató	Rendezési- és táblázat	Rendezési- és táblázat	Rendezési- és táblázat	EFO	Keresőszó- pos	Körforgó	Lancolatos	Szövegek öltetés és cím- tárgyszavak	WNU
Helyettesítés	.	V	V					V	.		.
Kapcsolás, Rendező- Kötő-
Szerpeljelölés											.
Elemzés	.	.									.
Rendezés	.	.									.
Súlyozás											.
Besorolás, Fogalomleíró- Dokumentumleíró- Vezérlő- Vezérlőfejezés	V	V	V	.	.	V	V	V	.	.	.
Betűrendes- Szakrendes- Számsorrendes	V	V	V	.	.	.	V
Egyszeres- Többszörös	V	V	V	.	.	V	V
Az általában végzett műveletek száma	8	9	6	6	7	7	6	6	7	4	9

- elvégzett művelet
- V vagyilagosan végzett művelet
- x az egymást kizáró vagyilagos műveletformák egy műveletnek számítanak.

1. táblázat

A jelentősebb visszakereső rendszerekkel általában végzett műveletek

feltárni mindazokat a belső törvényszerűségeket, amelyek egy nagyvolumenű ismeretanyag adatszerű részleteinek bármikor elvégezhető újra-fellelését, tehát folyamatos latens birtoklását lehetővé teszik. Ez az információ-visszakeresésnek a bevezetőben említett harmadik sajátossága, amelyik megkülönbözteti a bibliográfiától és az irodalomkutatótól.

A másik következtetés az információ-visszakeresés és a gépesítés kapcsolatára vonatkozik. A műveletek átfogó ismertetéséből is kitűnik, hogy napjainkban ezek zömét az ismeretanyag feltárója csak szellemi munkával végezheti el - és kevés a valószínűsége, hogy legtöbbjüket egyhamar gépesíteni lehessen. /Az információ-visszakeresésben főként a tárolási és a tényleges visszakeresési feladatok egyik része gépesíthető./ Ez pedig azt jelenti, hogy az információ-visszakeresésnek hasznos segédeszköze, sőt adott esetben az előfeltétele is lehet a gépesítés - de nem szükségképpen a velejárója. Egy kisebb tájékoztatási szerv az információ-visszakeresés feltárási és visszakeresési mélységét kisebb terjedelmű anyagban gépesítés nélkül is elérheti. De a nagyobb terjedelmű anyagban is "független" ebben az értelemben az információ-visszakeresés feltárási problémaköre a gépesítéstől. És ma már egyre inkább nyilvánvaló, hogy az információ-visszakeresés két fő problémaköre közül a gépesítés műszaki megoldásai egyértelműbbek és megnyugtatóbbak, mint a feltárás és a visszakeresés tájékoztatótudományi eredményei.

::
:: ::

TRODALOM

- /1/ LUHN, H.P.: Keyword-in-Context Index for technical literature /KWIC-index/. ASDD Report RC-127. International Business Machines Corp., Yorktown Heights, N.Y. 1959.aug.31. 6.p.
- /2/ TAUBE, M.: Studies in Coordinate Indexing. 1-4.k. 1953-57, Documentation Incorporated.
- /3/ Az alkotó elemzés részleteit lásd: HOVÁNYI G.: Ranganathan Kettőspont Osztályozása, a tizedes osztályozások és a dokumentáció. OMKDK Évkönyv-1961. Bp. 1962. OMKDK, 244.p. /p.147-185./.
- /4/ VICKERY, B.C.: On Retrieval System Theory. London, 1961.Butterworths, 159 p. /p.33-34./.
- /5/ Information retrieval - a solution. - HOLM, B.E. - Chemical Engng.Progress, 57.k. 6.sz. 1961. p.73-78.
- /6/ A logician's reaction to recent theorizing in information search systems. - BAR-HILLEL, Y. - American Documentation, 1957. 8.sz.
- /7/ BARHYDT, G.: Introduction to information processing activities at Western Reserve University. Az Information Retrieval in Action, Cleveland, 1963. WRU. 323.p. kötetből /p. 25-30./.
- /8/ PERRY, J.W.; KENT, A.; BERRY, M.M.: Machine literature searching. New York, London, 1956. Interscience Publ. 162 p.
- /9/ RANGANATHAN, S.R.: Elements of library classification. Ed. Palmer, B.I., London, 1959. Assos.Assist.Lib. 108.p.
- /10/ MARON, M.E.; KUHNS, J.L.; RAY, L.C.: Probabilistic indexing. Los Angeles, 1959. Woolridge, 127.p.
- /11/ A legismertebb: Automatic creation of literature abstracts. - LUHN, H.P. - IBM Journal of Research and Development, 2.k. 2.sz. 1958.ápr. p.159-165. - A kérdéssel a magyar nyelv szempontjából foglalkozó kiadvány: SZALAI S.: Gépi kivonatkesztés. Elméleti és kísérleti adalékok a magyar nyelvű tudományos közlemények gépi kivonatolásának kérdéséhez. Bp. 1963. OMKDK. 69.p. /Időszertű Műszaki Dokumentációs Kérdések 5./
- /12/ CLEVERDON, C.W.: Report on the testing and analysis of an investigation into the comparative efficiency of indexing systems. Cranfield, 1962. ASLIB. 305.p.
- /13/ BELZER, J.: Information retrieval on the GE 225 computer. Az Information Retrieval in Action, Cleveland, 1963. WRU. 323 p. kötetből /p. 31-39./.
- /14/ VICKERY, B.C.: Classification and indexing in science. London, 1959. Butterworths. 235.p. /p. 93./.

Г. Кобани:

Проблемы поиска информационного материала: Операции с понятиями по вскрытию темы документов

Поиск информации отличается от библиографии и исследования литературы тремя особенностями: шире, глубже и более систематизирован метод его вскрытия. Примером этого последнего служит серия взаимосвязанных операций, которая может быть проведена с понятиями по вскрытию документов. Отдельные операции: замена избранных по желанию предметных слов определенными заглавными словами, их подключение в соответствии с категориями или уровнем понятий, дополнение их отметками об их роли, операции по их анализу, предназначению, значению и занесения в архив. Семь главнейших операций и таблицы демонстрируют примеры наиболее распространенных систем поиска, чтобы показать на осуществление каких операций способны эти системы.

HOVÁNYI, G.: Probleme der Rechercharbeiten

Drei Kennzeichen unterscheiden die Informationsrecherche von den Bibliographien und von den Literatur-recherchen: eine breitere Basis des erschlossenen Gebietes, grössere Tiefe der Erschliessungsarbeit und eine systematischere Methode der Erschliessung. Ein Beispiel für letztere bildet die zusammenhängende Operationsreihe, welche mit den das Sachgebiet der Dokumente erschliessenden Begriffen durchgeführt werden kann. Die einzelnen Operationen sind die folgenden: Ersetzen der beliebig gewählten Stichwörter durch festgesetzte Schlagwörter, Kopplung derselben nach Kategorien oder Begriffs-niveaus, Ergänzung durch ihre Rolle angegebende Bezeichnungen, Analyse, Zuordnung, Festsetzung ihrer Bedeutung und schliesslich Einordnung in das Depot. Die sieben wichtigsten Operationen werden anhand von Beispielen aus den meist verbreiteten Recherche-Systemen gezeigt, und eine tabellarische Zusammenfassung gibt an, welche Systeme sich im allgemeinen für die Durchführung der einzelnen Operationen eignen.

+
++ ++