

Az Aberystwyth-i jelentés

HORVÁTH Tibor

Az E. Michael KEEN és Jeremy A. DIGGER által összeállított Report of an information science index languages test c. kiadvány alapján (Department of Information Retrieval Studies, College of Librarianship Wales, Aberystwyth. Aberystwyth, 1972. 3 Part) készült szemle.

A Könyvtári Figyelő 1968. évi 5. számában (351-369. p.) ismertette a cranfieldi kísérleteket és összegezte annak eredményeit. Mint ismeretes, Cranfieldben harminchárom osztályozási nyelvet, rendszert vetettek össze, konkluzióit, adatait azóta világszerte hivatkozzák. Az osztályozó rendszerek egybevetése, az értékelés mutatóinak kidolgozása ugyyszólván külön tudományos diszciplínát eredményezett az informatikán belül. Bár az értékelés szempont-rendszere eléggé kidolgozott, szinte valamennyi e témában publikáló szerző utal arra, hogy a tájékoztatástudomány, ezen belül az információkereső nyelvekkel foglalkozó tág témakör is inkább "puha" tudomány, szemben az egzakttságot vindikáló "kemény" tudományokkal. A konkluziók ezért mindig az adott feltételek keretei között érvényesek, nagyobb általánosításra csak fenntartásokkal lehet vállalkozni.

Ezeknek a fenntartásoknak elismerése mellett a Wales-i egyetem Könyvtártudományi Fakultás munkaközössége, E.M. KEEN és J.A. DIGGER vezetésével öt információkereső nyelvet, osztályozó rendszert hasonlított össze. A KDSZ már referálta M. Keen összefoglaló cikkét, (1973/985.) a teljes kutatási jelentés azonban most áll rendelkezésünkre.

Az öt osztályozási nyelv a következő.

1. Tömörített kifejezések osztályozó nyelve, posztkoordinációval. (Compressed Term language), tulajdonképpen tezaszuson nyugvó osztályozó nyelvet jelent.
2. Ellenőrzetlen osztályozó nyelv, magyarul talán a "szabad deszkriptorok" kifejezés a legmegfelelőbb. Posztkoordinált, egytagu, természetes nyelvű, kulcszavas osztályozást jelent.
3. Hierarchikusan strukturált nyelv posztkoordinációval.
4. Hierarchikusan strukturált prekoordinált osztályozó nyelv.
5. J. Farradane un. relacionális, hierarchikus, prekoordinált osztályozó rendszere (Relational language). Farradane kilenc un. operátort használ az osztályozó kifejezések közti kapcsolatok és hivatkozások jelölésére.

A dokumentumbázis nagysága 800 informatikai tárgyú mű, vagy annak referátuma volt, kivéve a relacionális osztályozást, ahol az osztályozott dokumentumok száma 241 volt. Az osztályozás egyrészt a teljes szöveg, másrészt a referátumok alapján ment végbe.

A vizsgálat feltételei közé tartozott, hogy az osztályozási nyelveket különböző részletességben alkalmazták és különbséget tettek a kifejezések specifikussági foka között is. Az osztályozás részletessége "(exhaustivity" - az osztályozás kimerítő volta)

azt mutatja, hogy a dokumentumot hány témakörre lehetett bontani az osztályozás során. (Egy terminológiai megjegyzés kívánkozik ide. Az informatikai szakszótár magyar változatának építésekor a "teljesség" műszót, amely az exhaustivity legjobb megfelelője, az angol "recall" - visszahívás magyar ekvivalensének foglaltuk le, ugyanis a recall mutató a visszakeresett releváns dokumentumok arányát mutatja az összes relevánshoz képest, tehát azt mutatja, hogy a releváns dokumentumok mennyire teljesen kerültek elő a visszakeresés során. Az osztályozás részletessége, az "exhaustivity" pedig annak mérőszáma, hogy az osztályozás során kimerítően irták-e le a témákat, azaz egy, kettő, ... stb. témára bontották-e. Nem azonos a részletesség az un. feltárt-sággal (angolban density vagy depth of indexing) amely azt mutatja, hány osztályozási kifejezés jut átlagosan egy-egy dokumentumra.)

Az Averystwyth-i kísérletekben az osztályozás részletességében (exhaustivity, tehát témákra bontásban) közepes (2 témára bontást) és magas (három témára bontást) vezettek be. A specifikusságnak három fokozatát különböztették meg: hierarchikus indexelést (1. fokozat) specifikus indexelés (2. fokozat) és redundáns indexelést (3. fokozat). Nyilvánvaló, hogy a hierarchikus indexelés a legkevésbé specifikus, ehhez képest dolgozik a többi rendszer bontottabb kifejezésekkel. Ezeknek a szempontoknak figyelembevételével az alábbi változatokat (ld. 1.sz. táblázat) állították elő. Az egybevetés nem úgy ment végbe, hogy minden szempontból valamennyi osztályozási rendszert minősítettek, hanem csak azokat a nyelveket, amelyek az adott szempontból egybevetethetők voltak, tehát amelyek éppen az egybevetés szempontjából különböztek egymástól. (1. sz. táblázat).

Végeredményben nyolc szempontu összehasonlítást végzett a kísérlet, mindegyikben egy-egy kérdéskört igyekeztek tisztázni. A nyolc összehasonlítás a következő. (Megadjuk azokat az indexelő nyelveket is, amelyekkel a konkrét egybevetést végezték.)

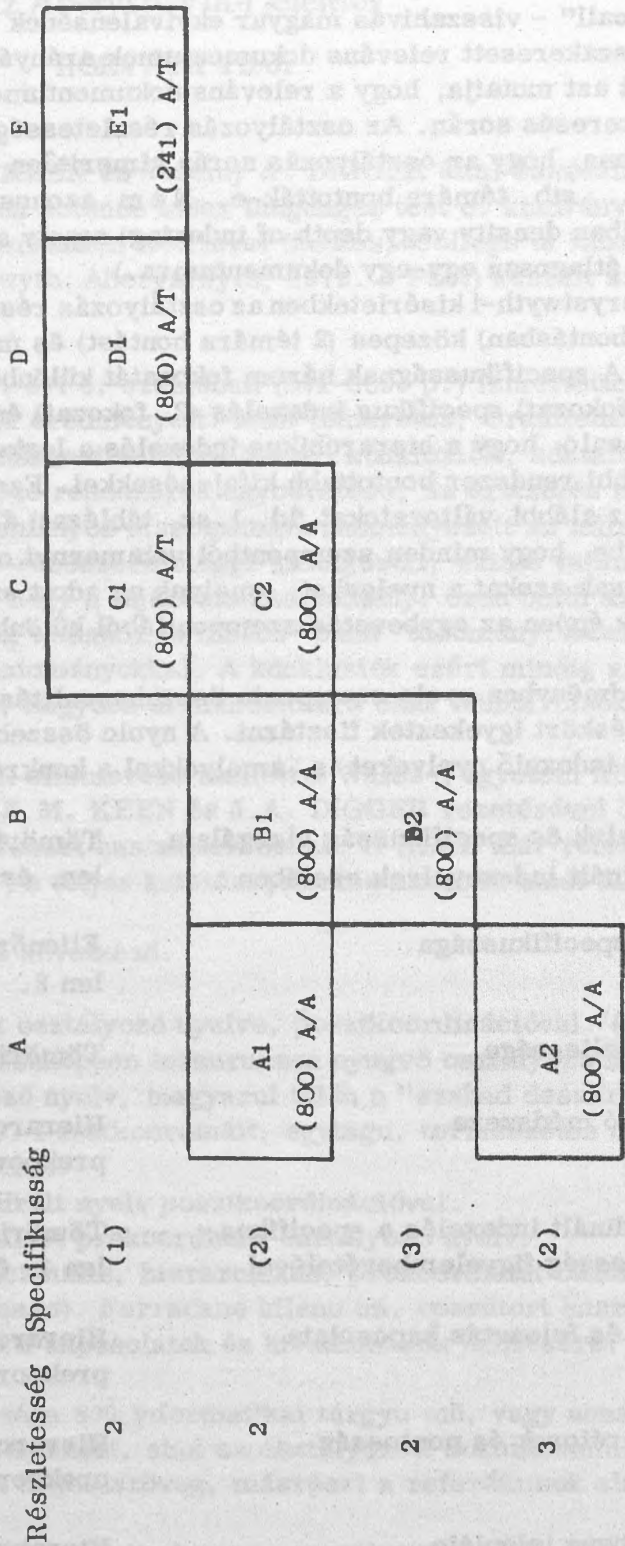
1. A kapcsolatok és specifikusság vizsgálata posztkoordinált indexnyelvek esetében	Tömörített kif. nyelve, Ellenőrizetlen, és Hierarch. posztk. 2.
2. Indexelés specifikussága	Ellenőrizetlen 1. és Ellenőrizetlen 2.
3. Indexelés teljessége	Tömörített 1. és Tömörített 2.
4. Koordináció módszere	Hierarch. posztk. 1. és Hierarch. prekoord.
5. Posztkoordinált indexelés a specifikusság és teljesség figyelembevételével	Tömörített kif. 1. és 2. Ellenőrizetlen 1. és 2. Hierarch. posztk. 2.
6. Pontosság és felosztás kapcsolata	Hierarch. posztk. 1. és Hierarch. prekoord.
7. Relációoperátorok és pontosság	Hierarch. prekoord. és Relacionális prekoord.
8. Eredeti szöveg jelenléte	Hierarch. posztk. 1. és Hierarch. posztk. 1. szöveggel

Prekoordinált

Posztkoordinált

Hierarchikus Hierarchikus Relacionális
posztkoord. prekoord.

Tömörített Ellenőrzetlen
kifejezés



A táblázat számai az egyes osztályozási nyelvek mellett a variánst jelentik és az első ábráról olvashatók le.

Az osztályozási kifejezések közti kapcsolatok mérésére a kapcsolatok számának és az indexelő kifejezések számának hányadosát használták. A "kapcsolatok" különböző utalásokban öltenek testet, mint a "lásd", "lásd még" utalók é.i.t. Ezek a kapcsolatok egyaránt lehetnek keresztutalások vagy különböző hierarchiaszintre való utalások, pl. az alá-, vagy fölérendeltség kifejezői.

A magyar szakirodalomban is ismeretes, - sőt, tananyagként is ismert - hogy a visszakereső rendszerek vizsgálatának legalapvetőbb mutatói a visszakeresés teljessége (recall) és pontossága. A teljességi mutató törtszámmal (vagy százalékban) azt fejezi ki, hogy a releváns dokumentumok hányada került elő. A pontosság azt mutatja, hogy a visszakeresett dokumentumok hányada releváns. A két mutató részletes kifejezésére itt nem térhetünk ki, remélve, hogy az olvasó előtt ismert fogalmakról van szó. Nem árt azonban összegezni az aberystwyth-i jelentés alapján - ők Vickeryre hivatkoznak ebben a témában -, hogy az osztályozási rendszerek különböző "fogásai" melyik mutatót befolyásolják. A 2. sz. táblázat ezt foglalja össze. Értelmezéséhez annyit lehet hozzáfűzni, hogy a "felosztás" ebben az esetben az osztályozási kifejezések csoportbasorolását jelenti. Ezért szerepel pl. a hierarchikus osztályozások prekoordinált változatánál. A kapcsolati jelölők pedig az összetett kifejezések hibás kapcsolatait vannak hivatva kiküszöbölni. (Pl. ha két tárgyszó: MÁGNES SZALAG és SZALAG KÁRTYA téves kapcsolással kiadja a MÁGNES KÁRTYA tárgyszót is.) A többi felsorolt elemzés, vagy műfogás a táblázat alapján megérthető. (2.sz. táblázat.)

Nincs mód arra, hogy a vizsgálat megszervezésének részleteit az olvasó elé tárjuk. Inkább érdekesek az eredmények. Mindjárt meg kell mondani, hogy az eredmények nem annyira az általánosítható tapasztalatokban, mint inkább a részletekben vannak. A részletekre jelen helyen viszont kevésbé lehet kitérni. Amit azonban az eddig elmondottakból érthetővé lehet tenni, azt a lehetőségekhez képest kiemeljük.

Az összegezés legáltalánosabb fokon végülis két fő konklúzióval zárult.

Az első az, hogy a vizsgált nyelvek nem mutattak szignifikáns különbségeket a visszakeresés hatékonyságában és hathatóságában, valójában nagy különbségek nem voltak észlelhetők.

A másik - már szembetűnőbb konklúzió - az volt, hogy az ellenőrizetlen kifejezésekkel végzett (szabad deszkriptoros) osztályozás mindazt nyújtani tudta, amit a tömörített osztályozási nyelv, tehát a kontrollált, tézauruszon nyugvó osztályozás. Az ellenőrizetlen osztályozás sohasem volt olyan rossz, mint a legrosszabb kontrollált, sem pedig olyan jó, mint a legjobb kontrollált, de egyetlen esetben sem voltak a különbségek szignifikánsak.

Részletesebben az alábbi konklúziók adódnak.

Specifikusság tekintetében az ellenőrizetlen nyelv és a hierarchikus nem nyújt eltérő pontosságot.

Magas teljességi követelmény mellett a koordinációs szintek redukálása illetve a kapcsolatok kiépítése nem eredményezte a többi nyelvtől való markáns eltérést.

A nyelvek specifikussági foka nincsen tisztán kapcsolatban a magas teljességi követelménnyel (recall), a magas specifikusság nem ad rosszabb teljességet, mint az alacsony.

Az osztályozási témákra való bontás a visszakeresésnél teljességében előnyöket jelent.

Az osztályozási témákra való bontás nem jelent különösebb előnyt vagy hátrányt a pontosságban.

	Tömörített ki- fejezések	Ellenőrzetlen	Hierarchikus		Relacionális
			Posztkoordinált	Prekoordinált	
Szemantikai Morfológiai Teljesség (recall)	Különböző írásmód	B	B	B	B
	Rövidítések	B	B	B	B
	Egysz. - többsz.	B	B	B	B
	Egyéb szóformák	B/1	L	b/L	L
	Ellentét	B		L	L
	Szinonima	B		B	B
	Generikus reláció	B		L	L
	Egyéb	B/1		L	L
	Homonima	P		P	P
	Koordináció	P	P	P	P
	Felosztás			P	P
Pontosság	Kapcsolatjeiőő	p	p	p	P
	Relációk	p	p	p	P

Kulcs:

	Teljesség (recall)		Pontosság
	Beépített	Kapcsolat	
	B	L	P
	b	l	p

Redundáns (ellenőrzetlen) osztályozás esetében a feltártság fokozása nem növeli a visszakeresés teljességét. Ugyanigy nem rontja és nem növeli a pontosságot, magas pontossági követelmények mellett.

Az alábbi táblázatok a felmérésnek mintegy összegezését adják. A 3. táblázat meg-
lehetősen zsufolt és bonyolult. Értelmezéséhez a függőleges oszlopok közül a "Wilcoxon"
feliratu szorul magyarázatra. A Wilcoxon próba a matematikai statisztikában két való-
színűségi változó rangsorolásán nyugszik. A két változó egy rangsorba kerül, ebben le-
het vizsgálni az egyik előfordulását. A vizsgálat a Wilcoxon próbát a nemreleváns doku-
mentumok vizsgálatánál alkalmazta. Az eredményt százalékban adja meg és az alacsony
százalékértékek azt jelentik, hogy nincs szignifikáns eltérés a két valószínűségi változó
között, jelen esetben a nemreleváns dokumentumok között.

E két táblázat eredményezte a fenti következtetéseket.

Ha az egész aberystwith-i kísérletet értékelni kellene, akkor azonnal szembetűnő,
hogy egész sor negatív megállapítást tartalmaz. Körülbelül oda jutottak el, hogy mi nem
befolyásolja a visszakeresés alapvető mutatóit, ti., a teljességi és pontossági mutató-
kat. A szerzők maguk írják, hogy nem derült ki világosan, melyik tényező játsza pl. a
meghatározó szerepet a teljességben. Ha az eredmény ilyen, akkor érdemes még
egyszer magának a vizsgálatnak módszereit felülvizsgálni. Kérdés ugyanis, hogy a ten-
ti következtetésen csak a vizsgálati feltételek között érvényesek, vagy a valóságot tük-
rözik, illetve hogy milyen mértékben tükrözik. Érdemes talán a szerzők által említett
egy példát citálni. A dokumentumok osztályozása ugy ment végbe, hogy az osztályozó
személy először hierarchikusan osztályozta a dokumentumot, aztán határozta meg a do-
kumentumhoz tartozó szabad deszkriptorokat. Kiderült, hogy a szabad tárgyszavak-ra
természetesen az osztályozónak azok a kifejezések jutottak eszébe, amelyekkel már
dolgozott az előző osztályozási rendszerben. Ilyen körülmények között nehéz a két nyelv
közti különbségeket megtalálni. A példa nem jellemzi a vizsgálatot, mert a hasonló ese-
teket kiszűrték s éppen az indexelő személyek magatartását is vizsgálni akarták, de
mindenestre utal arra, hogy a konkluziókba nagymértékben belejátszottak az alkalma-
zott vizsgálati feltételek is.

Az Aberystwith-i kísérlet ennek ellenére gazdagította ismereteinket arról az igen
nehéz kérdésköréről, hogy az osztályozási rendszer és az osztályozási fogások mi mó-
don befolyásolják a visszakeresés eredményeit. A probléma továbbra is nyitott.

ÖSSZEHASONLÍTÁS FAKTORAI	MAGAS PONTOSSÁGI KÖVETELMÉNY Közepes teljességi cél (közelítő)					
	Visszakeretett	Kérés különbségek	Kapott viszonylagos teljesség (recall)	Kérés különbségek	Nem releváns	Kérés különbségek
C1 Hierarchikus posztkoordinált Felosztás nélkül	6.20	10%	100%	3%	3.97	-
D1 Hierarchikus Felosztással	6.07		99%	-	3.87	8%
D1 Hierarchikus Felosztással, relációk nélkül	6.07	13%	100%	3%	3.87	-
E1 Relacionális Osztályozás Felosztással, relációkkal	5.82		98%	-	3.65	10%
C1 Hierarchikus, posztkoordinált Szöveg nélkül	8.98	63%	100%	33%	5.33	-
C1 Hierarchikus posztkoordinált Szöveggel	5.89		86,3%	-	2.87	56%

Rögzített keresés eredménye

ÖSSZEHASONLÍTÁS FAKTORAI			MAGAS PONTOSSÁGI KÖVETELMÉNY								MAGAS TELJESSÉGI KÖVETELMÉNY										
			Minimális teljességi cél Magasan releváns				Közepes teljességi cél Magasan releváns				Magas teljességi cél Magasan releváns				Magas teljességi cél Magasan és részlegesen releváns						
			Nem releváns különbségek				Nem releváns különbségek				Teljesség különbségek				Teljesség különbségek						
			Kapott teljes- ség (recall)	Nem releváns (cfd medián)	Kérések	Wilcoxon	Kapott teljesség Nem releváns (cfd medián)	Kérések	Wilcoxon	Másodlagos programok	Idő	Kapott teljesség	Kérések	Wilcoxon	Nem releváns (cfs medián)	Kapott teljesség	Kérések	Wilcoxon	Nem releváns (cfd medián)	Másodlagos programok	Idő
A ₁ Tömörített	szótár	273		T 32%		T 37%				T 19%	0,70 NS		T 32%	0,40 NS							
	specifikusság	20,5	98,4%	0,20	24%	97,3%	2,82	11%	2,95	9-06	87,1%	63%	13,1	74,7%	38%	18,9	6,33	18-39			
	kapcsolat	4,5		E 44%	1,42 NS	E 32%	0,73 NS			E 17%			E 30%								
B ₁ Ellenőrizetlen	szótár	1214		E 30%		E 37%				E 30%	1,05 NS		E 40%	1,16 NS							
	specifikusság	0,7	100%	0	28%	96,2%	0,92		4,49	9-14	83,6%	51%	8,9	70,4%	33%	13,8	8,17	19-16			
	kapcsolat	0,9		H 32%	0,07 NS	H 40%	0,70 NS			H 19%			H 27%								
C ₂ Hierarchikus	szótár	795		T 22%		T 25%				T 37%	2,25 Se 5%		T 46%	2,60 Se1%	10,7	8,98	19-55				
	specifikusság	2,8	100%	0	35%	98,2%	1,05	25%	5,13	11-55	80,9%	49%	6,4	68,3%	35%						
	kapcsolat	4,6		H 44%	1,54 NS	H 49%	1,83 NS			H 14%			H 19%								
Indexelés specifikussága	Specifikus	feltártság	7,2	100%	0	35%	96,2%	0,92	37%	4,49	9-14	83,6%	22%	8,93	70,4%	35%	13,8	8,17	19-16		
	2. típus (B1)				33%			25%				54%		27%							
	Redundáns	feltártság	10,5	100%	0	32%	97,9%	0,99	38%	0,29 NS	4,32	10-03	84,8%	24%	0,31 NS	8,59	71,4%	38%	0,09 NS	18,1	9,70
oszt. teljesség (exhaustivity)	Közepes	feltártság	6,9	98,4%	0,2	41%	97,3%	2,82	43%	0,42 NS	2,95	9-06	87,1%	13%	13,1	74,7%	19%	18,9	6,33	18-39	
	2. szint (A1)																				
	Magas	feltártság	10,7	100%	0,33	32%	100%	3,36	35%	3,35	8-23	92,3%	27%	1,63 NS	14,6	85,8%	56%	4,43 Se1%	24,4	6,51	17-45
Indexelő nyelv és szöveg	C1 Hierarchikus posztkoord. felosztás és szöveg nélkül		100%			98,2%				11-43	82,3%	38%	2,52 Se5%	66,9%	49%	2,8 Se5%			19-30		
												49%		29%							
	D1 Hierarchikus prekoord. felosztással és szöveggel	93,7%			90,7%					11-46	70,8%	13%		57,9%	22%				23-23		