

a szerkesztett kéziratot mikromásolat formájában, szelektív információszolgáltatás keretében terjesztik;

a szerző kéziratát annak tömörített változatával, következtetéseivel és kiemelt adataival együtt a kiadó adatbázisának megküldi, a felhasználó pedig megfelelően rendszerezett adatokat, következtetéseket stb. kap.

Modellek

Tökéletes információs rendszereket több mint 30 évi kutató és fejlesztő tevékenység után sem sikerült létrehozni. Az ilyen információs rendszerek mindmáig mutatkozó hiánya három modell segítségével magyarázható.

Az első modell egy tökéletesen, 100 százalékos eredménnyel működő dokumentumkereső rendszer. Egy ilyen rendszer nem oldaná meg a felhasználók túlterheltségének problémáját, hiszen még több releváns olvasnivalóval látná el őket. Semmi biztosíték sincs ugyanis arra, hogy erre a modellre a felhasználót segítő válogatás lenne jellemző, mint ahogyan ez a mai adatbázisokat sem jellemzi. Nem oldja meg, sőt súlyosbítja a problémát, ha a rendszer a felhasználót nem látja el az eredeti dokumentumokkal, hanem azokat csak jelzi. Az eredeti helyettesítő dokumentációs eszközök vagy bibliográfiai adatok szolgáltatása helyett információkat, adatokat szolgáltató rendszerre lenne szükség.

A második modell egy ugyancsak tökéletesen, 100 százalékos relevanciával működő, azonnali választ adó információs rendszer.

A körülményes dekódolást elkerülendő az output nyelve megegyezik a kérdező anyanyelvével. További követelmény a világos, teljes és tömör output. Ez az output azonban mindig a kérdés függvénye, hiszen a válasz éppen úgy lehet számadat mint egyszerű helyeslés, idézet, utalás vagy valamely módszer ismertetése.

A pontos válaszokat azonban nem lehet a gépbe táplálni, egyrészt, mert ehhez hiányzik a megfelelő számú emberi erő, másrészt, mert nem lehet minden lehetséges kérdésre előre felkészülni, a számítógép viszont mai fejlettségi szintjén még nem képes arra, hogy a primer és szekunder dokumentumokból álló inputból maga állítsa össze a megfelelő választ. Ez olyan korlátozó tényező, aminek következtében a felhasználónak be kell érnie a gépben tárolt információkkal, illetve újból csak válogatásra kényszerül. A helyzet a folyamatos tájékoztatásnál még nehezebb mint keresés esetén.

A két említett modell tehát nem hozza meg a kívánt megoldást, illetve nem valószínűsíthető meg. Marad a harmadik modell, melynek jellemzői a következők:

az output a felhasználó anyanyelvén, esetleg az általa ezen kívül legjobban értett természetes nyelven készül;

az output azonnal a felhasználóhoz kerül anélkül, hogy szükség lenne az eddigi keresési és továbbítási időre;

az állandóan naprakész output a felhasználó szűken körülhatárolt, speciális tárgykörére és meghatározott szempontjaira korlátozódik. Nem bibliográfia és nem tartalmaz felesleges anyagot;

az első output nem a felhasználót szavak tömegével elárasztó dokumentumokból, sem bibliográfiai adatokból vagy indexből, hanem az eredeti írásmű tized-századrészére tömörített dokumentációs feldolgozásból áll;

az output formája a különböző célokra, illetve a nyomtatott kötetből az ultramikrofilmig a felhasználás gyakoriságának megfelelően változik;

az outputot a felhasználó sorszám szerint rendezi; minden outputhoz külön nyomtatott index készül, melyet rendszeresen korszerűsítene;

a túlságosan kurrens információkat – pl. folyamatban lévő kutatásokról – on-line rendszer segítségével juttatják el a felhasználó munkahelyén vagy otthonában lévő adatvégállomásokra. Ezután ezek az információk is bekerülnek a fenti rendszerbe és a lehető legrövidebb időn belül a megfelelő formában is eljutnak a felhasználóhoz.

Ez a modell hatékony és technikailag is megvalósítható. Ennek, vagy hasonlóan működő rendszernek kiépítésével nem szabad tovább késlekedni.

Az első számú probléma megoldása után következik a második számú probléma, a kommunikáció hatékonysága. Akkor hatékony az olvasás, ha azt cselekvés, azaz az olvasott információt gyakorlati megvalósítás követi. Ez azonban az esetek túlnyomó többségében ma még nem történik meg.

Bár a valóban eredményes kutatásnak és fejlesztésnek nem csupán e két probléma a kerékkötője, de mindkettő súlyos és mindmáig megoldatlan.

/BERNIER, Ch. L.: Reading overload and cogency. = Information Processing and Management, 14. köt. 6. sz. 1978. p. 445–452./

(Dezső Zsigmondné)



A párhuzamos feldolgozások értékelése a dokumentumállományok feldolgozásakor

Az automatizált információs központok országos hálózatának létrehozása során fontossá vált az elsődleges dokumentumok feldolgozásának helyes koordinálása. A tematikai határterületek dokumentumait általában több információs szerv feldolgozza, és az információs kiadványok készítéséhez is gyakran ugyanazokat a forrásokat használják.

A párhuzamosság fokának megállapítása lehetővé teszi a dokumentumok ismételt feldolgozására fordított munka mennyiségének értékelését. A párhuzamosság fokának ismeretében lépéseket lehet tenni annak csökkentésére.

A párhuzamosság értékelésével egyelőre még keveset foglalkoztak. A Szovjetunióban mindössze néhány cikk jelent meg. Amerikában három nagy információs szolgálat – a *Biological Abstracts*, a *Chemical Abstracts Service* és az *Engineering Index* folytatott közös vizsgálatot, melynek során az egyes szolgálatok által feldolgozott állományokat dokumentumonként összehasonlították. Ez a módszer pontos eredményt ad, de nagyon munkaigényes.

A továbbiakban a cikk a párhuzamosság közvetett értékelési módszerét ismerteti. A módszer a források (a cikkeket tartalmazó folyóiratok) és a tematikai osztályok párhuzamosságának értékelésére épül. Ez a két adat minden esetben ismert, mivel a több információs központ által feldolgozott források és az azonos tematikai osztályok meghatározása a mágnesszalagos információcsere egyik előkészítő lépése. Nyilvánvaló a kapcsolat a dokumentumok és a tematikai osztályok ismétlődése között. Ugyanakkor az is kétségtelen, hogy megegyező tematikai osztályokkal dolgozó azonos források esetében a különböző információs szervek különböző cikkeket dolgozhatnak fel.

A közvetett módszer kipróbálására kísérletet folytattak, amelyet dokumentumonkénti összehasonlítással ellenőriztek. A kísérlethez két információs központ – a VINITI és a CNIITEI Priborosztroenija (Central'nij naucsno-iszszledovatel'szkij insztitut informacii i tehniko-ekonomicszeszkij iszszledovaniij priborosztroenija = Központi Műszeripari Tudományos–Műszaki és Gazdasági Kutatóintézet) – által egy év alatt feldolgozott számítástechnikai témájú állományt használtak.

A következő jelöléseket alkalmazták: \underline{A} a VINITI állományában, \underline{B} a CNIITEI Priborosztroenija állományában lévő dokumentumok halmaza. Ezeket a halmazokat forrásonként és tematikai osztályonként csoportosították a következő jelölésekkel:

\underline{a}_{ij} az \underline{A} halmazban az i jelű forráshoz és j jelű osztályba tartozó elem;

\underline{b}_{ij} a \underline{B} halmazban az i jelű forráshoz és a j jelű osztályba tartozó elem;

\underline{A}_{ij} az \underline{A} állományban az i jelű forrásból származó és a j jelű osztályba tartozó dokumentumok halmaza;

\underline{B}_{ij} a \underline{B} állományban az i jelű forrásból származó és a j jelű osztályba tartozó dokumentumok halmaza;

$|\underline{A}_{ij}|$ és $|\underline{B}_{ij}|$ a jelzett halmazok elemeinek száma.

A fentiekből következik, hogy:

- $\underline{a}_{ij} =$ 0, ha az \underline{A}_{ij} halmaz üres;
 1, ha az \underline{A}_{ij} halmaz nem üres;
- $\underline{b}_{ij} =$ 0, ha a \underline{B}_{ij} halmaz üres;
 1, ha a \underline{B}_{ij} halmaz nem üres.

Párhuzamosság tehát akkor van, ha az \underline{a}_{ij} és \underline{b}_{ij} elemek értéke nem nulla. Ezért a továbbiakban a $M^* = \{ \underline{a}_{ij} \}$ halmaz elemeit vizsgálták, amelyekre nézve teljesül az

$$\underline{a}_{ij} \cdot \underline{b}_{ij} = 1$$

feltétel.

Az \underline{A} és \underline{B} állományból párhuzamosan feldolgozott cikkek számát a következő képlettel határozták meg:

$$M_1 = \sum_{M^*} \min |\underline{A}_{ij}|, |\underline{B}_{ij}|.$$

Az összeget az M^* halmaz elemeiből képezik, az $|\underline{A}_{ij}|$ és $|\underline{B}_{ij}|$ minimális értékeiből. A mérőszám ilyen megválasztásának egyszerű oka van: a mindkét állományban meglévő, az i jelű forrásba és a j jelű osztályba tartozó dokumentumok száma biztosan nem több, mint az \underline{A}_{ij} és a \underline{B}_{ij} közül a kisebb.

A javasolt módszer ellenőrzésére M_1 -et M_2 -vel hasonlították össze, ami a dokumentumonkénti összehasonlításból adódott:

$$M_2 = |\underline{A} \cap \underline{B}|.$$

M_1 a párhuzamos feldolgozások számának felső értékét adja, azaz M_2 nem több, mint M_1 .

A kísérlethez a számítástechnika témakörét 64 osztályra osztották fel, mindkét állományt (\underline{A} -t és \underline{B} -t) a VINITI rubrikátora szerint osztályozták. A források száma 59 volt.

Nyilvánvalónak látszik, hogy az ismertetett módszer által kapott érték pontossága az osztályozási rendszer részletességétől függ. Ennek az igazolására a 64 osztályt nagyobb osztályokba fogták össze. Öt osztályozási szintet határoztak meg:

0 jelű szint	1 osztály
1 jelű szint	2 osztály
2 jelű szint	5 osztály
3 jelű szint	13 osztály
4 jelű szint	64 osztály

1. táblázat

A párhuzamos feldolgozások közvetett mutatójának függése az osztályozás mélységétől

Osztályozási szint	M_1	M_2	$\frac{M_2}{M_1}$
0	345	166	0,48
1	259	166	0,64
2	240	166	0,69
3	204	166	0,81
4	180	166	0,92

Az utolsó szint megfelel a VINITI rubrikátorának.

Az **A** állományban 7010, a **B**-ben 2485 dokumentum volt. A mindkét állományban feldolgozott forrásokból az **A**-ban szerepelt 934, a **B**-ben 373 cikk.

Az értékelés eredményét az 1. táblázat mutatja be.

A kísérlet eredményei igazolják, hogy az ezzel a módszerrel nyert értékek megfelelőek. Az értékelés pontossága az osztályozás mélységével valóban növekszik.

/KUZNECOV, B. A. – KALASNIKOV, É. V.: *Ocenka dublirovaniya pri obrabotke dokumental'nyh maszszivov. = Naucsno-Tehnicoszkaja Informacija, 2. sor. 4. sz. 1978. p. 18–19.*

(Viszocsekné Péteri Éva)



A Bradford-féle szóródás érvényesülése az információtudomány területén

A Bradford-féle szóródás a szélesen értelmezett információtudomány területén törvényszerűen érvényesül. Ha ezt megalapozottan számításba vennénk, akkor a hivatkozásvizsgálatok eredményeire támaszkodva, a jelenleginél sokkal megbízhatóbb és reprezentatívabb folyóirat-állománybázist lehetne az információtudomány művelői számára kialakítani egy-egy országban, így pl. Csehszlovákiában is.

Célszerű a hivatkozásvizsgálatok több szempontú elvégzése, ugyanis ennek alapján az állománygyarapítási koncepciót az egymásra vonatkoztatott eredményekre támaszkodva biztonságosabban alakíthatjuk ki. A szerzők a különböző szempontú vizsgálatokba 28 olyan információtudományi folyóiratot vontak be, amelyek a jegyzék összeállítására felkért szakértők szerint a tudományág legjelentősebb folyóiratai közé tartoznak.

Az információtudományt egészében átfogó *Information Science* két kötetében található összefoglaló jellegű tanulmányok és (legalább 30 hivatkozást tartalmazó) cikkek hivatkozásai az 1. táblázat szerint szóródnak a szóban forgó 28 folyóiratban.

A táblázatból megállapítható, hogy az információtudomány magja négy folyóiratban koncentrálódik, s további 11 tekinthető még e szempontból fontosnak. Sajnos, a csehszlovákiai könyvtárakban az említett folyóiratok egyike sem található meg.

Ha a szaktudomány egyes kiemelkedő témaköreire vonatkozó irodalom szóródását vizsgáljuk meg a 28 folyóiratban, akkor részben más eredményeket kapunk.

A mesterséges intelligenciával foglalkozó irodalom a 2. táblázat szerint oszlik el a folyóiratok között.

A formalizált nyelvek irodalmának szóródását a 3. táblázat mutatja (a mérés ez esetben Gladkij-

1. táblázat

Az információtudomány reprezentáltsága

Folyóirat	A hivatkozások száma
IEEE Transactions on Information Theory	25
Information and Control	24
IEEE Transactions on Automatic Control	15
Information Sciences	13
Avtomatika i Telemekhanika	12
ASME Transactions. Journal of Basic Engineering	10
Journal of Mathematical Analysis and Applications	8
SIAM Journal Control	6
Journal of ACM	6
Electronics Letters	6
Annals of Mathematical Statistics	6
Journal of Computer and Systems Sciences	5
Communications of ACM	5
IEEE Transactions on SSC	4
AIAA Journal	4
IRE Transactions	4
Journal of Medical Education	4
Journal of Optimization Theory and Application	3
Teorija Verovatnosztej i ee Primenenije	3
Proceedings of the IEEE	3
Lancet	3
BIT. Büro und Informationstechnik	3
Monatsschrift für Mathematik	3
Proceedings of the ACM	2
Automatica	2
Rozprawy Matematyczne	2
Problemü Peredacsi Informacii	2
International Journal of Control	2

2. táblázat

A mesterséges intelligencia reprezentáltsága

Folyóirat	A hivatkozások száma
IEEE Transactions on Information Theory	22
Proceedings of the International Joint Conference on Artificial Intelligence	20
Machine Intelligence	16
Artificial Intelligence	15
Communications of ACM	15
IEEE Transactions on Systems, Man, Cybernetics	15