

OSZTÁLYOZÁS

A fénylyukkártya,
mint a központi adattárolás eszköze

A fénylyukkártyának – a többi rendelkezésre álló lyukkártyarendszert is figyelembe véve – olyan előnyei vannak, amelyek különösen *alkalmassá teszik felhasználását központi tárolók esetében*. Központi tárolók létrehozása azoknál a tájékoztatási szerveknél szükséges, amelyeknek nemcsak egyetlen, hanem több – a legkülönfélébb kérdésekkel jelentkező – felhasználó igényeit is ki kell elégíteniük.

A központi tároló rendeltetése a visszakeresés útján történő információszolgáltatás. A visszakeresés minőségét – egyéb tényezők mellett – jelentős mértékben meghatározza a szolgáltatott információk száma. Ha ez a szám alacsony vagy esetleg nulla, a felhasználó elégedetlen: a minőség nem kielégítő. Az információk indokolatlanul magas száma sem hozza meg a kívánt eredményt.

A rendszer kialakítása előtt *célszerű összeállítani a kijelölt szakterület deskriptorainak jegyzékét*. Nem szükséges abszolút teljességre törekedni; olyan kifejezéseket nem kell felvenni, amelyek – tapasztalat szerint – ritkán fordulnak elő. Minél terjedelmesebb ugyanis a jegyzék, annál nehezebb a kezelése.

Abból kell kiindulni, hogy egy-egy fénylyukkártyán 7000 hely áll rendelkezésre; ezáltal 7000 dokumentum tárolására alkalmas. Egydimenziós tárolás esetén a deskriptor-jegyzék terjedelmének növekedésével egyenes arányban csökken a lyukasztások száma valamennyi fénylyukkártyán; a kártyák száma viszont ennek megfelelően emelkedik. (1. táblázat)

A teaurusz készítésénél nagy segítséget jelenthet a *Brooks-Carruthers szabály* (Statistical methods in meteorology, London, 1950) alkalmazása, amely kimondja, hogy annyi osztályt kell alkotni, amennyi az *n*-számú pozíció logaritmusának háromszorosának felel

Az egyes fénylyukkártyákon található lyukasztások száma a deskriptorok számához viszonyítva az egydimenziós tárolás esetén

1. táblázat

A deskriptor-jegyzék terjedelme	Deskriptorok száma/dokumentum	Összes lyukasztás száma	Lyukasztott pozíciók száma/fénylyukkártya	A deskriptor-jegyzék terjedelme	Deskriptorok száma/dokumentum	Összes lyukasztás száma	Lyukasztott pozíciók száma/fénylyukkártya
10	5	35 000	3500	10	10	70 000	7000
50	5	35 000	700	50	10	70 000	1400
100	5	35 000	350	100	10	70 000	700
1000	5	35 000	35	1000	10	70 000	70
2000	5	35 000	17	2000	10	70 000	35
3000	5	35 000	12	3000	10	70 000	23
4000	5	35 000	9	4000	10	70 000	18

Az egyes fénylyukkártyákon található lyukasztások száma a deskriptorok számához viszonyítva a többdimenziós tárolás esetén

2. táblázat

A deskriptor faja	A deskriptorok száma	Deskriptorok száma/dokumentum	Az összes lyukasztás száma	Lyukasztott pozíciók száma/fénylyukkártya	A deskriptor faja	A deskriptorok száma	Deskriptorok száma/dokumentum	Az összes lyukasztás száma	Lyukasztott pozíciók száma/fénylyukkártya
Vezérdeskriptor	10	3	21 000	2100	Vezérdeskriptor	10	2	14 000	1400
Fődeskriptor	100	3	21 000	210	Fődeskriptor	100	2	14 000	140
Deskriptor	1000	3	21 000	21	Deskriptor	1000	2	14 000	14
Vezérdeskriptor	20	3	21 000	1050	Vezérdeskriptor	20	2	14 000	700
Fődeskriptor	200	3	21 000	105	Fődeskriptor	200	2	14 000	70
Deskriptor	2000	3	21 000	10	Deskriptor	2000	2	14 000	2
Vezérdeskriptor	30	3	21 000	700	Vezérdeskriptor	30	2	14 000	466
Fődeskriptor	300	3	21 000	70	Fődeskriptor	300	2	14 000	47
Deskriptor	3000	3	21 000	7	Deskriptor	3000	2	14 000	5

meg. Ha tehát az $n = 7000$, akkor a legfelső szinten 12 ($3,8450 \times 3 = 11,54$), a második szinten 96; a harmadik szinten pedig 576 osztály kialakítása a célszerű. Ennek alapján a jegyzék 12 vezér-deszkriptort tartalmazna, amelyekhez 8 fődeszkriptor tartozik; ezek pedig 6 deszkriptort ölelnek fel. A jegyzék így 684, közel 700 egységet tartalmaz. *A kis dimenziójú jegyzék előnyös mind az indexelő, mind a tájékoztatási részleg számára; ezenkívül csökken a felhasználandó lyukkártyák száma is, ami anyagi vonatkozásban is előnyös (2. táblázat).*

Kisebbségi terjedelmű deszkriptor-jegyzék alkalmazása esetén a találatok valószínűsége lényegesen nagyobb, mint terjedelmesebb jegyzék használatakor (3. táblázat).

A deszkriptorok számának kívánatos csökkentése csak fokozottabb általánosítás (kevésbé mély indexelés) útján érhető el. A kifejezések sokaságából nehéz a legmegfelelőbbet kiválasztani; a legnagyobb nehézséget és legtöbb munkát azonban a megfelelő hierarchia kialakítása igényli. A befektetett munka megtérül, mivel ezáltal nemcsak a visszakeresés eredményességének esélyei javulnak, hanem könnyebbé válik a bejelölési művelet és az indexelés is.

3. táblázat

A találatok valószínűsége három fénylyukkártya kombinációja esetében különböző terjedelmű deszkriptorjegyzékekre vonatkozóan

A deszkriptorjegyzék terjedelme	Deszkriptorok száma/ dokumentum	Az összes lyukasztás száma	A lyukasztott pozíciók száma/ fénylyukkártya	A találatok valószínűsége
10	3	21 000	2100	0,900
25	3	21 000	840	0,360
50	3	21 000	420	0,180
100	3	21 000	210	0,090
250	3	21 000	84	0,036
500	3	21 000	42	0,018
1000	3	21 000	21	0,009
2000	3	21 000	10	0,004
3000	3	21 000	7	0,003

/SCHULZ, K.: Zur Anwendung der Sichtlochkartei = Informatik, 23. köt. 2. sz. 1976. p. 13-15./



(Takáchné Tóth Mária)

TÁJÉKOZTATÁSGÉPESÍTÉS

Adatbázisok és adatbázis-szolgáltatások értékelésének és kiválasztásának kritériumai az Egyesült Államokban

Az adatbázis alatt a továbbiakban *géppel olvasható információgyűjtemény* értendő, e kereten belül is elsősorban a bibliográfiai adatbázisokról esik szó.

A számítógépes adatbázisok használatát ma elsősorban az teszi szükségessé, hogy az információ-visszakeresés hagyományos, manuális módszerei többé nem vezethetnek a kívánt eredményre. Ennek megfelelően ma már gépi adatbázissal rendelkezik a természettudományok és műszaki tudományok nagy része, s néhány társadalomtudomány is. Ezek az adatbázisok négy csoportba sorolhatók:

szakterületiek (CAS Condensates, COMPENDEX stb.);

célra orientáltak (pl. NASA);

problémára orientáltak (szállítás, környezetvédelem stb.);

multidiszciplinárisak (pl. az ISI hivatkozási indexei).

Az USA-ban az adatbázisok többségét kezdetben állami szervek, intézmények hozták létre, s az első, adatbázisok feldolgozásán, szolgáltatásán alapuló központok kialakítását is az állam szorgalmazta. Ez a

tendencia időközben módosult, s *ma sokkal inkább a kereskedelmi jelleg érvényesül: az adatbázisok előállítói, a szolgáltató központok kereskedelmi jelleggel, önfenn-tartóként üzemelnek.*

Az adatbázisok használata 1974 óta ugrásszerűen növekedett. Csak egy példa: 1971-ben az USA-ban mintegy 50 ezer SDI (Selective Dissemination of Information = Szelektív információterjesztés) profil futott (a retrospektív keresések száma lényegesen kisebb lehetett). Jelenleg az on-line retrospektív keresések száma évente 700 ezer körül van.

Az adatbázisokkal kapcsolatban álló intézmények ma már szükségképpen hálózatokba szerveződnek: a munka mennyisége teszi ezt elkerülhetetlenné. *Ebből a folyamatból a könyvtárak sem vonhatják ki magukat.* Számos könyvtár szeretne bizonyos adatbázis-szolgáltatásokat nyújtani használói számára, de nem tudja, hogyan fogjon hozzá. Hogyan értékelje, válassza ki a megfelelő adatbázisokat és szolgáltatásokat?

* * *

Az adatbázis-szolgáltatásokat, igénybevételük formája szerint, *batch üzemmódú* vagy *on-line* szolgáltatásokra szokás osztani. Lényegesebb szempont azonban e szolgáltatások tartalma és keresési módja. A három fő típus: