

SZÁMÍTÓGÉPES INFORMÁCIÓKERESÉS

A nyilvánosan hozzáférhető adatbázisok fejlődése Nyugat-Európában

Az Eusidic Database Guide a „kereskedelmi úton hozzáférhető” kifejezést a „nyilvánosan hozzáférhető” kifejezéssel azonos értelemben használja.

Az Eusidic* két egymástól alapvetően különböző adatbázissal foglalkozik:

1. *a bibliográfiai fájlokkal*, amelyek egy-egy dokumentumot ajánlanak a felhasználó figyelmébe, aki aztán valamelyik könyvtárban „manuálisan” megkeresheti a dokumentumot, abban a reményben, hogy megtalálja benne a keresett információt. Ebbe a kategóriába tartoznak azok az adatbázisok is, amelyek információforrásként egy-egy szervezetet vagy személyt jelölnek meg, pl. az USA Világkereskedelmi Információs Központjának fájlljai.
2. *az adatbankokkal*, amelyek a végső, tényleges információt nyújtják. Rendszerint numerikus információ-ról van szó, a két fő tartalmi típust a közgazdasági és demográfiai statisztikák, illetve a vegyületek fizikai tulajdonságaira vonatkozó adatok reprezentálják. Sajátos típust képviselnek a jogi szövegek adatbankjai és a terminológiai adatbankok.

Az adatbázisok kialakulása

Az adatbázisok széles körű létrehozása a második világháború utáni információs és számítógépes fellendülés idejére esik. Az egyesek által információrobbanásnak nevezett jelenség három fő okra vezethető vissza: a háborús idők okozta hatalmas lemaradás, a multinacionalizálódás, különös tekintettel az Egyesült Államokra, valamint az új, gyakran interdiszciplináris feladatok.

A bibliográfiai indexeket számítógép segítségével egyre jobb hatásokkal sikerült előállítani. Számuk ennek következtében exponenciálisan nőtt. A gépi adathordozón levő melléktermékeket még offline üzemmódban is sokkal gyorsabban lehetett végigkeresni, mint manuálisan. Hát még interaktív módon! Óriási megtakarítást jelentett, hogy táv-lekérdezés révén nagy fenntartási költségű fájlokon sok felhasználó osztozhatott.

Ami a numerikus adatokat illeti, csak fokozatosan derült ki, hogy számos ilyen adatállomány jelentős érdeklődésre tarthat számot. Az adatbankok anyagát aztán egyre fejlettebb software-rel együtt kezdték közre-

adni, vagyis az eredményeket egyszersmind alá lehetett vetni aritmetikai feldolgozásnak is, további modellek építésére, grafikus megjelenítésekre is lehetőség nyílt.

A hatvanas évek elejétől a hetvenes évek elejéig

Kétségbenvonhatatlan, hogy a fellendülésben az állami indíttatású vállalkozásoknak volt elsődleges szerepük. Az 1. táblázatból látható, hogy először azokhoz a nagy indexekhez lehetett online hozzáférni, amelyek révén az Egyesült Államok kormánya a háború után megnövekedett információs igényeknek igyekezett megfelelni az űrkutatási és nukleáris technika, a nevelés- és egészségügy, majd a környezeti problémák, valamint a világ élelmiszer- és energiaellátása területén. Ezek ma is a világ legnépszerűbb bibliográfiai fájlljai közé tartoznak: 1978-ban a világ valamennyi online bibliográfiai keresésének egy negyede a MEDLARS-ot érintette.

Más ágazatok fájlljai, tipikusan a tudományos társulatok irodalmi indexei, eleinte csak az egyes nagy cégek belső használatára szolgáltak, utóbb aztán sok rendszer csatlakozott az állam által indított online szolgáltatásokhoz.

A hetvenes évek elejétől napjainkig

Az 1. táblázat azt is kimutatja, hogy míg a kormány valamint a különféle egyesületek és alapítványok főként a tudományos-műszaki folyóirat- és report-irodalomra koncentráltak, addig a *profitra épülő szektor* már kezdetben is más anyagok feltárására fordította figyelmét, úm.

- *a szabadalmakra*, elsőként a gyógyszerészet körében (pl. a Derwent Szabadalmi Világ-indexe, először a Farmdoc, amelynek testvérfájlla volt a Ringdoc a folyóiratcikkekéről),
- *az üzletre és közigazgatásra* (pl. a CPI, az International Systems Design időközben megszűnt vegyi marketing fájlla és a ma is virágzó Congressional Information Service Index),
- *a humán tudományokra* (pl. az ABC-Clio Amerika történelméről és az amerikai életről, jelentős mértékben a Xerox cég automatikus bibliográfiai szolgáltatásai).

A fenti törekvések realizálása közben újfajta keresési módok is megjelentek:

* European Association of Scientific Information Dissemination Centres = Európai Tudományos Információs Központok Szövetsége.

A mágnesszalagon és online hozzáférhető bibliográfiai
fájlok alakulása 1963 és 1970 között*

1963–1964	1965–1966	1967–1968	1969–1970	1971–1972	1973–1974	1975–1976	1977–1978
ÁLLAMI SZEKTOR							
NASA	online						
	ERIC		online				
		MEDLARS	online				
		ENDS		online			
		NSA	online				
			NTIS online				
			CAIN		online		
			TOXLINE	online			
NEM PROFITRA ÉPÜLŐ SZEKTOR							
API					online		
	METADEX				online		
		GeoRef			online		
		CACon		online			
		Embase			online		
			COMPENDEX		online		
			BIOSIS		online		
			INSPEC		online		
			SPIN				online
PROFITRA ÉPÜLŐ SZEKTOR							
SCI					online		
Ringdoc						online	
WPI						online	
	ABS						
	AHL						online
		ICRS					
			GeoArchive				online
			CPI				
			CIS/Index				online

* A cikk táblázatainak adatai egy 1979. jan. 1-i Eusidic-felmérés eredményét tükrözik.

- a kémiai jelölés szerinti (pl. az ISI [Institute for Scientific Information = Tudományos Információs Intézet, USA] Index Chemicus regisztrációs rendszer),
- a hivatkozási indexek családja szerinti (pl. az ISI Science Citation Index szalagjai).

Ez a nem műszaki–tudományos anyagokra való hangsúlyeltolódás voltaképpen a hetvenes évek tematikai forradalmává lett (2. táblázat).

Financiális szempontból és a jövő szempontjából még forradalmibb változás volt a hetvenes években, hogy a hangsúly az adatbázisokon belül az adatbankokra helyeződött át. A 3. táblázat jól illusztrálja ezt a trendet (noha a felölelt anyaggal és a fogalmak értelmezésével kapcsolatos bizonytalanságok folytán nem egészen rigorózus a statisztika).

Az USA (tudomásunk szerint) 337 adatbankjának csaknem kétharmada két nagy funkciókörhöz kapcsolódik, az egyiket a kormány „szabályozó misszió”-jának, a

2 táblázat

A bibliográfiai fájlok megoszlása
tárgykörök szerint

	1972		1978	
	db	%	db	%
Természettudomány	45	46	173	31
Műszaki tudomány	24	25	135	24
Egyéb	17	17	179	32
Általános	12	12	69	13

másikat a profitra dolgozó szektor „ökonometrikus misszió”-jának nevezhetjük. Az első kategóriába 67 adatbank tartozik, köztük a környezetvédelmi hivatal 25, a nemzeti egészségügy 14, a nemzeti szabványügyi iroda 8, az élelmezés- és gyógyszerügy 7 adatbankja; a másik kategóriát 117 adatbank példázza (az összesen 146 üzleti—közgazdasági profilú közül), többek között a Chase 23, a Data Resources Inc. 17 és a Standard & Poor's 8 adatbankkal.

Az adatbankok sikere fenomenális. Egyetlen angol üzleti adatbank (a dataSTREAM International) csaknem háromszor annyi online keresést regisztrál, mint az összes bibliográfiai fájl együttvéve. A 4. táblázat tisztán online vonatkozásban mutatja be, hogyan nőtt az (első-sorban amerikai üzleti) adatbankok aránya a bibliográfiai fájlokkal szemben.

Nem forradalmi, de jelentős a harmadik, a földrajzi jellegű eltolódás is: a bibliográfiai fájlok dominanciája az Egyesült Államokból Európába kerül át. Ennek négy fő oka van:

1. Az Egyesült Királyság nagy nyelvi előnyt élvez, az angol nyelv a huszadik század latinja, az angol indexek a világon mindenütt eladhatók, nem is beszélve a magas fejlettségű angol nyelvű piacról.
2. A többi európai ország nyelvileg hátrányban van. Ez is készítette a francia, szovjet és nyugatnémet kormányt, hogy hazai változatokat szubvencionáljanak (pl. PASCAL, Referativnij Zsurnal, Internationale Dokumentationsgesellschaft für Chemie).
3. Igen sok kormányközi szervezet székhelye Európába települt, pl. a FAO, ILO, UNESCO stb.
4. Az amerikai adatbázisok nem ölelik fel teljes mértékben az európai és nem-amerikai anyagot.

Az 5. táblázatból látható, hogy 1979 elején az egyik legnépszerűbb amerikai cégtől beszerezhető bibliográfiai fájloknak csaknem a fele Európából származik.

Európában a fájlok növekedési üteme is nagyobb. Részletes adatokat tartalmaz a 6. táblázat, amelyből az egyes szektorok részvállalása is megállapítható: a világ legnagyobb bibliográfiai fájllai közül az állami szektorhoz tartozó 16 fájl évi 4 millió 140 ezer rekorddal gyarapodik, a nem profitra dolgozó szektor 5 fájlja évi 800 ezer rekorddal, a profitra dolgozó szektor 14 fájlja pedig évi 2 millió 640 ezer rekorddal.

Hogyan tovább?

Európából lényegében hiányoznak azok az adatbázisok, amelyek nem tudományos—műszaki, hanem ún. „egyéb” információkat tartalmaznak és általában regionális jellegűek, valamint nyilvánvaló, hogy több meglévő európai adatbankot kell online hozzáférhetővé tenni. E

3. táblázat

A bibliográfiai fájlok és az adatbankok számának
alakulása 1972 és 1978 között

	ASIS		Európai adatbázisok			Eusidic Database Guide	1979. I. 1-i felmérés
	1972	1975	1976	1977	1978	1979	
Bibliográfiai fájlok	98	335	337	422	533	556	
Adatbankok	—	51	149	268	568	609	
Összesen	98	386	486	690	1101	1165	

Online bibliográfiai fájlok és adatbankok eredetük és kereskedelmi terjesztők száma szerint

Kereskedelmi terjesztők száma	Bibliográfiai fájlok		Adatbankok	
	Európaiak	Európán kívüliek	Európaiak	Európán kívüliek
0	190	117	187	115
1	79	79	28	211
2	33	25	7	38
3	2	20	2	6
4		3		9
5	INSPEC	BIOSIS CACon MEDLARS NTIS Pollution Abs.		IFS Teleprice 70 WPI/CPI
6				Value Line
7				Census
:				
:				
15				NBER

A Lockheed Information Systems bibliográfiai fájljai

	65-66	67-68	69-70	71-72	73-74	75-76	77-78	79
Összesen	1	1	4	6	18	63	92	110
Ebből európai	0	0	0	0	3	30	39	54

két lemaradással szemben van viszont Európában két izgalmas műszaki újdonság: az angol posta *Prestel kísérlete*, amely az adatbázisok eddig széles skálájának nyithat eddig példátlanul tág piacot, a *European Space Agency* pedig 1980-ban munkába állíthatja az első európai nem-kísérleti kommunikációs műholdat, ami az információterjesztés költség- és hatékonyság-mutatóit jelentősen megjavíthatja.

A legjobb eredményeket bizonyára a kormány, a nem profitra épülő szektor és az üzleti világ folyamatos további együttműködése révén lehet elérni, olyan információs közösségben, amelyet egyaránt egyesítenek az

ambiciózus szakmai célok és a szabad információáramlás társadalmi célja. Nemcsak adatbázisokat kell importálni az Egyesült Államokból, hanem az információ szabadságára vonatkozó törvényt is, hogy felszabaduljon számos, a kormányok által jelenleg visszatartott adatbank.

A jövőben valószínűleg olyan feladatokat kell találni, amilyenek kezdetben is kiváltották az adatbázisok fennállítását. Így a munkanélküliség, a szabadidő eltöltése, a kelet-nyugati kapcsolatok, a harmadik világgal való kapcsolatok izgalmas és izgató problémái adják talán majd a nyugat-európai adatbázisok egy részének tárgyát.

A világ legnagyobb bibliográfiai „termőterületei”

	Európában	Európán kívül
Évi 1 599 000 – 800 000 rekorddal bővülő fájlok		
INPADOC – szabadalmak	1 000 000	
Évi 799 000 – 400 000 rekorddal bővülő fájlok		
IDC-Patentdatenbank – szabadalmak	700 000	
SCI szalagok – műsz. és term.tud.		520 000
PASCAL – műsz. és term.tud.	500 000	
Referativnij Zsurnal – műsz. és term.tud.	500 000	
WPI – szabadalmak	500 000	
CÁ – vegyészet		420 000
JICST File – műsz. és term.tud.		420 000
	<u>2 200 000</u>	<u>1 360 000</u>
Évi 399 000 – 200 000 rekorddal bővülő fájlok		
BIOSIS Previews – biológia		250 000
Embase – orvostudomány	250 000	
MEDLARS – orvostudomány		250 000
Information Bank – hírlapi információ		200 000
	<u>250 000</u>	<u>700 000</u>
Évi 199 000 – 100 000 rekorddal bővülő fájlok		
CAB – mezőgazdaság	150 000	
(CT, a CA része		150 000)
F and S Indexes – üzleti		150 000
LIBRIS – könyvek	150 000	
AGRICOLA – mezőgazdaság		140 000
Federal Index – USA, politikai		140 000
(PASCAL vegyészet, a PASCAL része	130 000)	
(PASCAL orvostudomány, a PASCAL része	130 000)	
SSCI szalagok – „puha” tudományok		130 000
(CAIN, az AGRICOLA része		120 000)
CPI – konferenciák anyagai		120 000
EDB – energia		120 000
Ringdoc – farmakológia	120 000	
IRL adatbázis – az élettel kapcs. tud., „life sciences”	110 000	
MARC (Könyvek)		110 000
(PASCAL biológia, a PASCAL része	110 000)	
Registre des Testaments – végrendeletek	110 000	
AGRIS – mezőgazdaság	100 000	
INSPEC A – fizika	100 000	
GeoArchive – geo-tudományok	100 000	
Libcon/E – könyvek		100 000
Wall Street Journal Newspaper Index		100 000
Washington Post Newspaper Index		100 000
	<u>940 000</u>	<u>1 210 000</u>
Összesen	4 390 000	3 270 000

/TOMBERG, A.: *The development of commercially available databases in Europe. = Online Review, 3. köt. 4. sz. 1979. p. 343–352./*

(Szöllősy Éva)