

Az index jelenleg mintegy 3000 profilt tartalmaz. A profil Ontario állami intézményeiben, egyetemeken, minisztériumaiban és egyéb szerveinél működő szakemberek, kisebb munkacsoportok, részlegek adatait esetleg egy szervezet teljes kutató tevékenységét jelenti.

Az index mintegy 30 000 tárgyszava és ugyanannyi kereszthivatkozása egy vagy több profilszámra utal, ami két jegyzék segítségével vezet a kért információhoz. Az egyik a profilok jegyzéke – célja, hogy a kapott kifejezés relevanciája a teljes szöveggörnyezetben legyen ellenőrizhető –, míg a másik a szakemberek nevét, telefonszámát, címét, valamint a kommunikációs feltételeket (pl. nyelv, egyes megkötések) közli.



## ADATBÁZISOK INFORMÁCIÓSZOLGÁLTATÁSOK

### Az USA adatbázisainak fejlődése 1966–1975 között

#### Bevezetés

Az elmúlt tíz évben rohamosan növekedett a gépi adatbázisok száma, mérete, valamint a felhasználók tömege. Ezzel párhuzamosan az adatbázisok tartalmilag is jelentősen bővültek mind a feldolgozott forrásokat, mind az átfogott időszakot és tárgyköröket illetően.

Ez a fejlődés egyúttal az adatbázisok működtetési költségeinek, valamint az információkeresés fajlagos költségének és árának csökkenésével, ill. az adatbázisok finanszírozásának változásával, az adatbázisokat használó szervezetek körének jelentős bővülésével.

Mind több közlemény jelent meg az adatbázisokról, s növekedett az adatbázisokat előállító feldolgozó és felhasználó intézmények száma is. A kereséshez használt szótárak fejlődése egységesítési törekvéseket váltott ki.

Az adatbázisokkal kapcsolatos rendszerszervezési munka olyan színvonalra fejlődött, hogy az adatbázisok hálózattá szervezésének kérdéseivel is eredményesen foglalkoznak.

#### A beszerezhető gépi adatbázisok száma és mérete

Míg 1965-ben még kevesebb mint 20, addig 1975-ben már több mint 300 adattár volt forgalomban, közülük mintegy 50 már on-line rendszerben is hozzáférhető volt. Míg tíz évvel ezelőtt egy-egy adatbázis évente átlag néhány ezer dokumentumról szóló információval gyarapodott, addig ma már olyanok is léteznek, amelyeknek évi gyarapodása több százezer (pl. a CAS és az ISI évi gyarapodása 400 ezer körül van, a BIOSIS-é pedig 240 ezer).

Nyomatásban az index nem jelenik meg, minden kérés a CISTI szűrőjén megy keresztül. Nevek közléséhez a CISTI köteles az érintett személyek, ill. intézmények engedélyét előzetesen beszerezni. Természetesen a kérdeztettek az adatok közlését meg is tagadhatják.

Az új vállalkozás egyelőre még kevésbé ismert, a CISTI most kezdi széles körű népszerűsítését. Amennyiben a szolgáltatás iránt nagyobb érdeklődés mutatkozik, azt Kanada egész területére kiterjesztik.

*[Infoscope, 1977. 6. sz. p. 9–10.]*

(D. É.)

#### A felhasználók és a keresések száma

1965-ben még csak mintegy 10 ezerre, 1970-ben pedig már legalább 100 ezerre becsülhető a gépi adatbázisok használóinak száma. Az Egyesült Államokban és Kanadában 1975-ben csak az on-line retrospektív irodalomkutatások száma megközelítette az egymilliót.

#### A felölelt szakterületek

A kezdetben elsősorban kormányzati szervezetekben (NASA, AEC, NLM) jöttek létre adatbázisok. Később a különböző referáló szolgáltatásokra alapozott gépi adatbázisok kialakulásával megnövekedett a mérnöki és alkalmazott tudományos információk aránya a feldolgozásban. Végül megjelentek a társadalomtudományi gépi adatbázisok is. Újabban az amerikai kutatás fejlesztésének lassulásával párhuzamosan megnövekedett a kereskedelmi és gazdasági információk beépítése az adatbázisokba. Legújabbban már olyan adatbázisokat is hoznak létre, amelyeknek tartalma nem a kutatót, fejlesztőt vagy az üzletembert, hanem az egyszerű állampolgárt érdekli (pl. fogyasztás, jog, gyermeknevelés, gyógyszerek stb.).

#### Retrospektív adatbázisok

A tíz évvel ezelőtti gépi adatbázisok csupán rövid idővel korábban keletkezett információkat tartalmaztak és így retrospektív szolgáltatásokról értelmetlenség lett volna beszélni. Ma már vannak adatbázisok, amelyek több mint tíz évre visszamenőleg tárolnak információkat.

#### A feldolgozott dokumentumok fajtái

Folyóiratcikkek, kutatási jelentések, szabadalmi dokumentumok mellett fokozatosan megkezdődött a mo-

nografikus kiadványok, valamint a nem időtálló dokumentumok feldolgozása és a bibliográfiai jellegű adatbázisok mellett az *adatinformációkat* tartalmazó adatbázisok létrehozása. A távközlési hálózatok elterjedésével, valamint a tároló kapacitás olcsóbbodásával lassan megvalósul mind a politikai, mind a tudományos-műszaki elektronikus újság gondolata. Adatbázisokat hoztak létre a folyamatban lévő kutatásokról, ilyenek pl. az SSIE, a CRIS, a Register of Research Projects, a CANCERPROJ és a HRIS.

Határozott irányzatot lehet kimutatni a *numerikus adatbázisok* létrehozatalában is. 1975-ben megalakult az *Association of Public Data Users* azzal a céllal, hogy egyesítse az elsősorban a kormányzati szerveknél előállított adatok iránt érdeklődő felhasználókat. A Predicasts és a Lockheed közös erőfeszítésével kialakított és működtetett PATS adatbázisban – amely a DIALOG távközlési hálózaton keresztül hozzáférhető – már 100 ezer statisztikai idősor található és kezelhető (átalakítható, összefüggéseiben vizsgálható), a felhasználó igényeinek megfelelően.

#### Pénzügyi források

Az adatbázisok tárgyköreinek kiszélesedésével párhuzamosan nagymértékben megváltoztak az adatbázisok létrehozására és működtetésére felhasznált pénzügyi források. Egyes adatbázisokat a kormány finanszírozott. A referáló és indexelő szolgáltatásokban keletkezett – tulajdonképpen a nyomtatott kiadványok szövegtükrének mágnesszalaggal vezérelt fényszedéséhez használt – mágnesszalagos adatbázisok terjesztésének sikere után *üzleti vállalkozásként is állítanak elő és működtetnek adatbázisokat*, például az ISI, a Predicasts, az Excerpta Medica Foundation, a Chase Manhattan Bank stb.

#### A szolgáltatások ára

A kormányzati adatbázisok ingyen álltak a kormányhivatalok rendelkezésére, sőt később ezek és az egyéb kormányzati segítséggel kifejlesztett adatbázisok nyilvános használata is ingyenes volt. A támogatást azonban később fokozatosan csökkentették, sok esetben meg is vonták az adatbázisoktól azzal az indokkal, hogy a kormány csak az adatbázisok létrejöttét óhajtotta ösztönözni. Ezzel egyidőben *üzleti vállalkozások is létrejöttek*. Olyan tendencia alakult ki, amely szerint a szolgáltatásoknak ára van, amit a felhasználónak meg kell fizetnie és ebben az árban benne kell lennie a ráfordításokon kívül a nyereségnek is. Létrejöttek a profitra dolgozó információs ipar vállalatai (ISI, Predicasts, Inc., Lockheed Information Systems, SDC). E fejlődési irányzatot azonban csak a legkorszerűbb és egyre olcsóbbá váló technika alkalmazása teszi lehetővé.

Míg 1965-ben a keresés valamely egymillió tételt tartalmazó adattárban profilonként kb. 1000 dollárba került, addig 1970-ben az ára 100 dollár, 1975-ben pedig mintegy 10–50 dollár.

#### Az intézményi és egyéni felhasználók

A korábban egyedüli felhasználóként szereplő kormányzati szervek mellett megjelentek felhasználóként az iparvállalatok, az egyetemek, kutatóintézetek, sőt a nyilvános könyvtárak is.

*Megváltozott a felhasználás módja, és változtak maguk a felhasználók is.* A saját célra történő adatbázis létrehozása és felhasználása mellett elterjedt az adatbázisok eladása és emellett a külső felhasználók részére történő keresés is. Így megjelentek a vásárolt adatbázisokat feldolgozó és az azokból ismét külsők részére szolgáltató, ma már elterjedten on-line üzemmódban, távközlési hálózat igénybevételével dolgozó *gépi adatbázisok*.

#### Szakmai szervezetek

Az információs ipar eddig körvonalazott fejlődése odavezetett, hogy a géppel olvasható adatbázisokkal különféle szakmai testületek kezdtek foglalkozni, illetve kifejezetten ebben a tárgykörben új szakmai szervezetek jöttek létre.

Az ASIS keretében létrehoztak a szelektív információterjesztéssel, illetve az adatbázisok on-line, interaktív felhasználásával foglalkozó csoportokat (SIG/SDI és SIG/UOI). 1968-ban alakult meg az ASIDIC és 1970-ben a hasonló célú európai szervezet, az EUSIDIC. Az IIA sok tagja szintén érdekelt a gépi adatbázisokban.

Végül 1975-ben az ASIDIC kifejezésre juttatta, hogy egyaránt foglalkozni óhajt az adattárak előállításával, feldolgozásával (az adatbázisok működtetésével) és az adatbázis-szolgáltatások közvetítőivel is. Ezért az ASIDIC rövidítés feloldását a következőképpen változtatta meg: *Association of Information and Dissemination Centers*.

#### Az adatbázisokról szóló szakirodalom

Jelenleg évente közel 500 kutatási jelentés és folyóiratcikk foglalkozik a témával. A *Bulletin of the ASIS* állandó rovatot nyitott az adatbázisok tárgykörének. Az Association of Computing Machinery (USA) 1975-ben indította meg a *Data Base* c. folyóiratot. A már említett szakmai szervezetek és az adatbázisokat előállító és működtető vállalatok mind adnak ki híreket tartalmazó folyóiratokat. Ezekon kívül több monográfia, szemle, bibliográfia, jegyzék is megjelent az adatbázisokról és

felhasználásukról. Az ASIS évente megjelenő szemlékötete az *Annual Review of Information Science and Technology* is egyre több fejezetében foglalkozik a témával.

### Szabványosítás

A korai, belső használatú adattárak virágzása idején nem volt szükség szabványosításra. A későbbi fejlődés hozta magával az *ANSI Z 39.2-1971* jelű (szabvány mágnesszalagos bibliográfiai információcserére) amerikai szabványt és alkalmazását először a Kongresszusi Könyvtár MARC rendszerében. Azóta ezt a szabványt sok adatbázis-előállító és üzemeltető információs központ fogadta el.

A feldolgozó központoknál azonban rendszerint saját formátumra konvertálják a beérkező mágnesszalagokat, így tulajdonképpen nem lényeges, hogy azok milyen szabvány szerint készültek (nem beszélve most az adatelemekről, csak az adatbázis formátumáról). Ha azonban egy központban sok adatbázist egyidejűleg dolgoznának fel és azok azonos formátumban érkeznének, programozási és számítógépi idő megtakarítható lenne. Az NSF, ASIDIC, EUSIDIC, NFAIS és ICSU/AB 1975 májusi értekezletén ezért javaslatot készített a bibliográfiai adattárak adatelemeinek egységes minimális készletére.

### Az adatbázisok szervezése

A kezdeti, elsősorban nem a keresés (hanem például a fényszedés) igényét kielégítő adatbázisok szervezési módja helyett egyre inkább figyelembe veszik azt, hogy az adatbázisokat ma már *döntő mértékben információkeresésre használják*. Olyan törzset (master file) alakítanak ki és hoznak ezért forgalomba, amelyből az igényeknek megfelelő különböző formátumú és tartalmú végtermékek állíthatók elő.

### On-line keresési software

Tíz évvel ezelőtt még csak elképzelés volt, kormányzati segítséggel azonban azóta magasszintű rendszereket fejlesztettek ki. Ilyenek például a RECON III és a DIALOG (a Lockheed rendszerben), illetve az ELHILL és az ORBIT (az SDC rendszerében).

### Szótárak adatbázisokban történő kereséshez

A beszerezhető adatbázisok sokfélesége oda vezetett, hogy ma az irányzat inkább előnyben részesíti a címekben, kulcsszavakban, kulcsszavas mondatokban és referá-

tumokban való szabad keresést a magas szinten strukturált információkereső nyelvek kidolgozásával és használatával szemben.

### Hálózatba szervezés

Jelenleg ez a fejlesztési irány mutatkozik a legerőteljesebbnek, mert lehetővé teszi a földrajzilag kis területen ki nem használható nagy, retrospektív adatbázisok létrehozását és felhasználását, azaz sok felhasználónak sok adatbázisból *kooperációban történő on-line üzemmódu* kiszolgálását.

### A fejlődés és problémái

Az információs központok vezetésében kialakuló költségszemlélet és költségtudatos gazdálkodás jó irányban befolyásolta a fejlődést. Helyes, hogy a felhasználók a szolgáltatást megfizetik. A kormányzat segítsége az adatbázisok kifejlesztésében, valamint a jobb irányítás és a korszerű technológia alkalmazása lehetővé tette, hogy *a szolgáltatások csökkenő árát a felhasználók meg is tudják fizetni*. A további kutatás-fejlesztés azonban már csak kormányzati segítséggel mehet végbe. Ugyanakkor az adatbázisok működtetésének aránylag nagy költségei csak a szolgáltatások további tökéletesítésével és mind szélesebb körű használatával csökkenthetők.

Ezzel szemben negatív hatást gyakorol az információs iparra a kormányzati és a magánszektor egyre változó kölcsönös kapcsolata. Így például a kormányzat, amely korábban segítette az információs ipar fejlődését, az utóbbi időben néhány szolgáltatását már nem bocsátja a nyilvánosság rendelkezésére. Egyes esetekben a kormányzat által létrehozott adatbázis csak kormányzati szervek útján férhető hozzá, de ingyenesen vagy csupán minimális díjért. Az alacsony árú, dotált kormányzati szolgáltatások ismét előtérbe kerülnek, s ezzel az információs iparnak jelentős versenytársa alakul ki.

Problémát jelent az is, hogy a hálózatok megszervezésével *csökken egy-egy referáló és indexelő szolgáltatás előfizetőinek a száma* mind a mágnesszalagos, mind a nyomtatott formátumra. Ez viszont csak az árak növelésével ellensúlyozható. Számos további probléma is jelentkezik. Így például a szerzői jog kérdése; az adatok tulajdonjoga; az adatbázisok és a software szabadalmaztatása; az információk export–import tarifáinak szabályozása stb.

Nem kevés a gond az informatika *oktatásával* kapcsolatban sem: a tantervek ugyanis nem tartottak lépést az adatbázisok nemzetközi szinten is tapasztalható fejlődésével.

Az eredeti dokumentumokkal kapcsolatos legnagyobb probléma az, hogy az USA-ban nincs olyan koordinált hálózat vagy központ, amely esetleg on-line

módon meg tudná adni valamely dokumentum lelőhelyét.

#### Következtetések

Az adatbázisok rohamosan fejlődő világában jelentkező problémák megoldása csak *együttműködéssel* lehetséges. Így például az input költségek megosztása érdekében már létrejött együttműködés elsősorban a felesleges párhuzamos feldolgozások megszüntetésére a CAS és a BIOSIS, az AIP és az Ei stb. között. Az együttműködési lehetőségek kiszélesítése érdekében az ASIDIC, EUSIDIC, NFAIS és az ICSU/AB közösen *foglalkoznak az adat-elemek egységesítésével*.

A költségek jelenleginél nagyobb mértékű megosztását célozza az NSF Office of Science Information Research és az Office of Computing Activities közös kutatása a hálózatszervezés továbbfejlesztésére.

A MIT fejlesztési munkájának célja a hálózati, on-line, interaktív rendszerekhez való hozzáférés megkönnyítése egységes és egyszerűen kezelhető gépi nyelv kialakításával. Az adatbázisok közötti tájékozódást, valamint az adott kérdésnek megfelelő adatbázisok kiválasztását segítené elő a terminál útján hozzáférhető *Data Base Selector*, a legközelebb működő megfelelő adatbázis megtalálását pedig a szintén terminállal elérhető *Data Base Mapping* rendszer. Mindkét rendszert az Illinois-i egyetemen dolgozták ki az NSF anyagi támogatásával.

Az áttekintést 34 tételes hivatkozásjegyzék egészíti ki.

#### Rövidítések jegyzéke

AEC	Atomic Energy Commission
AIP	American Institute of Physics
ANSI	American National Standards Institute
ASIDIC	Association of Scientific Information Dissemination Centers
ASIS	American Society for Information Science
BIOSIS	Bioscience Information Service
CAS	Chemical Abstracts Service
CRIS	Current Research Information System
EUSIDIC	European Association of Scientific Information Dissemination Centers
HRIS	Highway Research Information Service
ICSU/AB	International Council of Scientific Unions Abstracting Board
IIA	Information Industry Association
ISI	Institut for Scientific Information
MARC	Machine Readable Cataloging
MIT	Massachusetts Institute of Technology

NASA	National Aeronautic and Space Administration
NFAIS	National Federation of Abstracting and Indexing Services
NLM	National Library of Medicine
NSF	National Science Foundation
PATS	Predicasts Abstract Terminal System
RECON III	Remote Console
SDC	System Development Corporation
SIG/SDI	Special Interest Groups on Selective Dissemination of Information
SIG/UOI	Special Interest Groups on User On-line Interaction
SSIE	Smithsonian Science Information Exchange

*/WILLIAMS, M. E.: Data bases – A history of developments and trends from 1966 through 1975 = Journal of the American Society for Information Science, 28. köt. 2. sz. 1977. p. 71–78./*

(Schiff Ervin)



#### Szelektív információterjesztés a wroclawi Műszaki Egyetemen

A wroclawi Műszaki Egyetemen először 1971-ben foglalkoztak a bibliográfiai-információs folyamatok automatizálásával. Eredeti elképzelésük szerint saját feltáró munkájukra, majd később az ágazati információs központoknál folyó feltárára alapozva óhajtottak kiépíteni információs szolgáltatást. Mindkét elképzelés életképtelen volt a feltárás „hézagos” volta és az egyöntetű munkához szükséges szabványok és szabályzatok hiánya miatt.

Így kezdtek el foglalkozni a nemzetközi mágnesszalagos szolgáltatásokkal és 80 ilyen szolgáltatás elemzése alapján állapították meg, hogy az *INSPEC (Information Services in Physics, Electrotechnology, Computers and Control = Fizikai, Elektrotechnikai, Elektronikai, Számítástechnikai Információs Szolgálat)* és a *CAC (Chemical Abstracts Condensates)* a Műszaki Egyetem dolgozóinak tematikai információs igényeit 60%-ban ki tudja elégíteni. A két mágnesszalagot 1974 közepétől, az *ISMEC (Information Services in Mechanical Engineering = Gépészeti Tájékoztató Szolgálat)* és a *PASCAL (Programme Appliqué à la Sélection et à la Compilation Automatiques de la Litterature = Automatikus Irodalomgyűjtő és Válogató Program)* szalagokat – amelyek kiegészítik az előzőeket – 1975-től szerzik be.

Az e szalagokra alapozott *szelektív információterjesztés* hamar népszerű lett: 1974 második felében 131,