

jelei — mégsem az elektronikus számítógépeké lesz a kulcsszerep.

Véleményünk szerint a jövő szempontjából a döntő tényező a tudósok és szakemberek információigénye, a felhasználók igényeinek növekedése és egyre bonyolultabbá válása lesz. Emellett azt is figyelembe kell vennünk, hogy jelentős mértékben kiszélesedik az információ előállítóinak és felhasználóinak társadalmi összetétele.

Elsősorban tehát ezek a folyamatok fogják meghatározni a tudományos információ szférájában bekövetkező változások lényegét, belső tartalmát.

A tudományos információ fejlődésével kapcsolatos korábbi prognózisokban sok volt a tévedés, mivel nem vették figyelembe a tudományos kommunikáció törvényszerűségeit, a tudományos információ specifikus tulajdonságait.

Megfigyeléseink alapján a prognosztizálás során a következő sajátosságokat kell szem előtt tartani:

1. A tudományos kommunikációban — ellentétben más társadalmi folyamatokkal és jelenségekkel — a hagyományos eszközöket és módszereket nem szorítják ki az újak, hanem funkcióikat átcsoportosítva a régiek és az újak közös rendszerbe szerveződnek.
2. A tudományos kommunikáció általunk ismert törvényszerűségei statisztikai jellegűek, és bennük a különböző vizsgált jelenségek expo-

nenciális eloszlást mutatnak. Ez a körülmény arra enged következtetni, hogy létezik a kommunikációban egy általános törvény, de ezt eddig még nem sikerült általános érvényű formában kifejezni. Ezt feltétlenül tudnunk kell, amikor a prognosztizálásban már megismert törvényszerűségeket figyelembe vesszük.

3. Az elkövetkezendő évtized fejlődési perspektívájának meghatározásakor arra a feltevésre kell támaszkodnunk, hogy a tudományos információ előállítóinak szociális bázisa jelentősen kiszélesedik, és a paradigmaváltás nemcsak a tudományos ismeretek tartalmában, hanem a tudományos kommunikáció módszereiben, formáiban és eszközeiben is bekövetkezik. Ahhoz, hogy az információs folyamatokat irányítani tudjuk, ismerünk kell ezeknek a változásoknak a jellegét és irányát, figyelemmel kell kísérnünk az informatikában folyó kutatásokat, és rendszeresen kell prognosztikai kutatásokat végezni.

*/MIHAJLOV, A.I.: O budušem naučnoj informacii = NTI, 1.Ser. 1985, 1. sz. p. 1–3./*

(Környei Márta)

## **A szakértői információs rendszerek kialakításának jelentősége az informatikában\***

Napjainkban az informatikai kutatásokban központi helyet foglalnak el az információkeresés elméletével és az információs rendszerek kialakításának elvi kérdéseivel foglalkozó kutatások. A céljuk az, hogy az emberi tudás megismerésében és feltárásában új, egyre hatékonyabb módszerek álljanak a rendelkezésünkre.

A XX. század közepéig a könyvtári (szak-, tárgy- és szerzői) katalógusok voltak a dokumentumokban felhalmozódott információk feltárásának egyedüli eszközei. A század ötvenes éveitől kezdve a koordinált indexelési módszer kidolgozása új utat

nyitott mind az információkeresésben, mind az információs folyamatok automatizálása területén. Ezt a módszert alkalmazzák napjainkban a legtöbb ún. deskriptoros információkereső rendszerben, melyekben a keresési folyamatok automatizálása a második és harmadik generációs számítógépek segítségével történik.

A keresőrendszerek fejlesztésén kívül elméleti kutatások folynak olyan logikai rendszerek kimunkálására, amelyek létrejöttével valóra válhatna az információk automatikus feldolgozása, valamint azoknak az információknak a "kinyerése", amelyeket csak rejtve tartalmaz a tudományos közlemények szövege. Ezeket a kutatásokat jelentős mértékben támogatja a számítástechnikai gyorsütemű fejlődése és az ötödik generációs számítógépek kifejlesztésének előrehaladása. Az ötödik generációs számítógépek legfőbb célkitűzése, hogy az adat- és

\* A szakértői rendszerekről ld. még: NOWAK, E.J. — SZABŁOWSKI, B.P.: Szakértői rendszerek a tudományos és műszaki információ terjesztésében (TMT, 32. köt. 7. sz. 1985. p. 338–341.) és YAGHAMAI, N.S. — MAXIN, J.A.: Szakértői rendszerek (TMT, 32. köt. 8–9. sz. 1985. p. 418–423.) — (Ref.: Szöllösy Éva).

információfeldolgozáson túlmenően a tudás- (ismeret-)feldolgozást is lehetővé tegyék. Ennek előfeltétele, hogy megismerjük és megértsük az ismeretek belső struktúráját, az adatok, tények, elméletek, feltevések közti összefüggéseket. Az ezzel kapcsolatos kutatások a matematikai logika, a logikai szemantika, a strukturális nyelvészet és egy sor más tudományterület eredményeire támaszkodnak.

### Adatbázis, adatbank

Könyvtárosok, információs szakemberek körében divattá vált manapság adatbázisokról, adatbankokról beszélni, ami arról tanuskodik, hogy a számítógép alkalmazása bekerült a szakmai köztudatba. A baj a fenti fogalmak értelmezésével van. Hasonló értelmezési zavar volt például, amikor az információkeresés elméletének fejlődésével a katalógusokat elkezdtek információkereső rendszernek nevezni — ami általánosságban igaz is —, miközben a katalógusok hagyományos szervezése mitsem változott. Ugyanígy: a mágnesszalagos információs kiadványokat is szeretik adatbázisnak nevezni — és ez sem teljesen helytelen értelmezés —, de az adatbázis fogalmának lényegét nem tükrözi.

Az *adatbázis* olyan számítógéppel támogatott speciális rendszer, amely a különböző feladatok megoldásához szükséges adatokat egy helyen gyűjti össze és biztosítja az adatok függetlenségét a feldolgozási programoktól. A korábban szinonimaként használt adatbank és adatbázis fogalmak jelentése különvált.

Az *adatbank* magában foglalja az adatbázisrendszert, valamint mindazokat a programokat, nyelvi, szervezési és technikai eszközöket, amelyek biztosítják az adatok központi tárolását és kollektív felhasználását.

Az adatbázisok létrejötte megteremtette a lehetőséget, hogy az automatikus információkereső rendszerek mellett kialakuljanak azok a logikai rendszerek, amelyek egyaránt képesek az információk elemzésére és szintézisére. Ez azt jelenti, hogy a rendszer képes megkönnyíteni a kutatás és a kutatók számára a meglévő és szükséges információk rendezését, egységesítését, statisztikai információk előállítását, az adatok párhuzamos sorai közötti kapcsolatok feltárását és egyéb "intellektuális" feladatok elvégzését.

### Intellektuális információs rendszerek — az információs rendszerek "intellektualizálása"

Az információs rendszerek továbbfejlesztésének útja a meglévő rendszerek "intellektualizálása",

melynek lényege, hogy a rendszer képes legyen:

- ◆ az adatok (ismeretek) nagy tömegének fontosság szerinti rendezésére,
- ◆ ebből a rendezett adathalmazból — a következtetés logikai eszközeivel — kivonni az összes szükséges adatot (ismeretet),
- ◆ értékelni, hogy a tárolt és kivont adatok (ismeretek) igazak, hamisak, bizonytalanok vagy értelmetlenek-e (ami egy metanyelv segítségével történik),
- ◆ a "külvilágból" kapott új információk nyomán új típusú kérdés-válasz párokat alkotni, valamint kiválasztani azokat a "nemtriviális" kérdéseket, amelyekre a rendszer válaszolni tud.

Az adat- illetve ismeretfeldolgozás során az empirikus összefüggések feltárására az információs rendszer a logika deduktív és induktív eljárásait alkalmazza.

Az adatbázis előbbi definíciójának analógiájára az *ismeretbázison* az egy-egy tudományterülethez vagy ismeretkörhöz tartozó igaz állítások számítógépes memóriában strukturáltan tárolt összességét értjük.

Ebben az összefüggésben az *adaton* a kiindulási adatokat értjük, melyek lehetnek term-ek, term-sorozatok, segédszimbólumok stb., az ismereten pedig azokat az adatokat, amelyek szerkezete a közlések közti kapcsolatokat adja meg. Az ismeretbázis megköveteli a kiindulási adatok (predikátumok) közötti kapcsolatok bizonyos mértékű axiomatizálását.

Az "intellektuális" információs rendszerek közül — ide tartoznak az információkereső rendszerek kivül a döntéshozatali, a diagnosztikai, az automatikus kérdés-válasz rendszerek — az ún. szakértői információs rendszerek kidolgozása nyújt komoly segítséget azokban az informatikai kutatásokban melyek az információs rendszerek és folyamatok intellektualizálásával foglalkoznak. A szakértői információs rendszerek lényege abban áll, hogy megfelelő logikai apparátus segítségével a kvalifikált szakértők gondolkodását, szellemi tevékenységét utánozzák. A rendszerek egyrészt bizonyos logikai-matematikai modelleket, másrészt a matematikai logika nyelvéhez hasonló metanyelvet használnak. Rendkívül fontos, hogy mivel az emberi gondolkodást akarják utánozni, olyan logikai modelleket válasszanak, amelyek a hiányos (nem teljes körű) információ esetében is biztosítják a feladat megoldását. Ilyen logikai modell az adatbázisszervezésből ismert relációs modell, melyben az adatábrázolás a predikátum-kapcsolatok rendszere segítségével valósul meg. Az eddig elmondottakon kívül e rendszerek működésében még ki kell dolgozni a kérdés-válasz predikátumok elméletét is a hiányos információk esetére.

## Az absztrakt információs rendszer és a szakértői információs rendszer

V.K. Finn nyomán az absztrakt információs rendszer alapvető paraméterei a következők:

1. az adatábrázolás nyelve — L (tárgynyelv)
2. az adatfeldolgozás nyelve — ML (metanyelv); a metanyelv a tárgynyelven ábrázolt kiindulási adatok kezelési, valamint a kérdés-válasz predikátumok nyelve,
3. a kiindulási adatokkal (predikátumokkal) kapcsolatos nemlogikai axiomák halmaza — S,
4. a megengedett kérdések halmaza — Q, ezek a metanyelv (ML) specifikus (kérdő) term-jei,
5. a kérdés-válasz predikátumok halmaza — K, mely a Q kérdéseire a kérdés-válasz eljárásokat tartalmazza metanyelven,
6. a lehetséges tömbök halmaza — amelyből minden egyes M tömb az L, ML, S, Q, K paraméterekkel együtt az információs rendszer állapotát [M] adja meg:

$$[M] = \langle L, ML, S, Q, K \rangle$$

A fentiek alapján az információs rendszert tekintjük úgy, mint a logikai struktúra  $\mathcal{R} = \langle L, ML, S, Q, K \rangle$  összes lehetséges állapotának  $\Theta$  [M] halmazát.

A szakértői információs rendszerekben Q a feladatok halmaza, S az ismeretek halmaza, K pedig azoknak a logikai és számítási eljárásoknak a halmaza, amelynek segítségével a rendszer képes utánozni a szakértő intellektuális tevékenységét.

A szakértői információs rendszereknek két típusát különböztetjük meg, az egyik típus a magyarázó, a másik a bemutató rendszer.

*Magyarázó* a szakértői információs rendszer akkor, ha az S halmaz axiomái és az M tömb megadott kezdeti állapota alapján bizonyos empirikus tényeket képes megmagyarázni:

$$(S \cup M) - \phi$$

ahol  $\phi$  — a magyarázott tények leírása; — — a levelezhetőségi reláció jele;  $\cup$  — pedig a halmazelméleti unió jele.

*Bemutató* a szakértői információs rendszer akkor, ha képes reprodukálni azt az utat is, aminek eredményeképpen megkapta a felállított hipotézisre a végkövetkeztetést (hipotézis-“demonstráció”).

A szakértői információs rendszerek továbbfejlesztésében az induktív eljárások eredményeként kapott hipotézisek automatikus előállítására területén folyó kutatások jelentenek komoly perspektívát.

### A JSM módszer

Az Összszövetségi Tudományos és Műszaki Információs Központban (Vsesoúznyj Institut Nauč-

noj i Techničeskoj Informacii; VINITI) 1979 óta folynak az automatikus hipotézis-előállítás módszerével kapcsolatos kutatások. A módszer kidolgozásában a valószínűségi következtetési eljárások szabályrendszerét alkalmazzák, amely John Stuart Mill angol filozófus indukciós logikai rendszerén nyugszik, az ő tiszteletére JSM módszernek nevezik.

A kutatások eredményeképpen kidolgoztak egy olyan hiányos információkkal rendelkező szakértői információs rendszert, amely a valószínűségi következtetés szabályainak segítségével képes hipotéziseket felállítani a kiindulási adatokból a hiányos információs szituációkra vonatkozóan.

### A rendszer működése

1. Az induktív következtetés szabályainak segítségével a rendszer hipotéziseket állít fel a kiinduló események közti korreláció meglétének vagy hiányának okáról (1. szabály).
2. Ezután az analógia alapján történő következtetési szabályok segítségével a megtalált ok(ok)at “átteszi” az adatbázisban levő bizonytalansági esetekre (2. szabály). Ez a bizonytalanság (hiányos információ) abból adódik, hogy nincs tudomásunk arról, hogy az E esemény rendelkezik-e A tulajdonsággal vagy tulajdonsághalmazzal.
3. A rendszer megállapítja, hogy  $E_1$  esemény része az E-nek, és  $E_1$  az A tulajdonság meglétének az oka (az 1. szabály segítségével). Ugyanakkor az adatbázisban nincs információ arról, hogy E rendelkezik-e A tulajdonsággal. Az X események, amelyek nem tartalmazzák az  $E_1$ -t, nem rendelkeznek A tulajdonsággal. Ezek után az analógia alapján történő következtetési szabályok szerint E esemény rendelkezik az A tulajdonsághalmazzal.
4. Ezután az adatbázis új állapotot vesz fel, amelyben a bizonytalanságot — ti., hogy az E esemény rendelkezik-e A tulajdonsággal — felváltja egy “pozitív” hipotézis, nevezetesen, hogy az E esemény bizonyos mértékű valószínűséggel rendelkezik A tulajdonsághalmazzal. Hasonló módon felállítanak egy “negatív” hipotézist is, miszerint az E esemény nem rendelkezik A tulajdonsággal. A hiányos információval rendelkező szakértői információs rendszer a következő feladatokat oldja meg automatikusan:
  - ◆ hipotéziseket állít fel a tulajdonságok meglétének, ill. hiányának okairól,
  - ◆ hipotéziseket állít fel a megfelelő tulajdonságokkal rendelkező eseményeknél a meglét okairól,
  - ◆ hipotéziseket állít fel (ezek induktív általánosítások) az empirikus összefüggésekről, mivel az

okok lehetnek összetett események, elemi okok Boole-algebrai kombinációi,

- ◆ rangsorolja a hipotéziseket a valószínűség mértéke alapján és attól függően, hogy milyen szabályok szerint állították fel őket, illetve, hogy az egyes szabályok alkalmazása során mennyi volt a lépések és az adatok száma,
- ◆ megállapítja azokat az eseteket, amikor az eseménytulajdonság párokra vonatkozóan két "konkurrens" hipotézis azonos valószínűséggel rendelkezik (adott esemény rendelkezik, ill. nem rendelkezik A tulajdonsággal),
- ◆ kiszámítja a kapott eredményt, és ennek alapján osztályozza az objektumokat.

A rendszerrel végzett eddigi kísérletek egytől-egyetig kedvezőek. Az adatbázisban szereplenie kell mind az események (objektumok), mind az eseményeket jellemző tulajdonságok halmazának. Az ese-

mények és tulajdonságok között létezniük kell empirikus összefüggéseknek. Emellett az adatbázisnak rejtve tartalmaznia kell bizonyos információt a tulajdonság- és eseményhalmaz közti korreláció meglétének, ill. hiányának okairól. Mindezeket a halmazokon meg kell adni azt a részlegesen (nem mindenütt) meghatározott kapcsolatot, amely az adatbázisban levő információ hiányos voltát tükrözi.

A kifejlesztett rendszer minden gyengén formalizált diszciplínában jól alkalmazható, vagyis olyan tudományterületen, ahol hiányzik az elméleti állítások, tételek szisztematikus axiomatizálása.

/GILÁREVSKIJ, R.S.: O značenii razrabotki ékspertnyh informacionnyh sistem v informatike = NTI, 2.Ser. 1984. 11.sz. p. 1-4./

(Környei Márta)

#### Az online és a hagyományos irodalomkeresés összevetése: a keresési stratégiát és a keresés hatásfokát vizsgáló kísérlet érdekes eredményei

Tíz évvel ezelőtt, amikor az online információkeresés még újdonság volt, az informatikai szakirodalom kedvencei közé tartozott egy téma gépi és kézi keresése eredményeinek összehasonlítása. A gépi módszerek azóta sokat fejlődtek, kiépültek a nagy, nemzetközi elérésű adatbázisok, így újra érdemes egy ilyen összehasonlítást elvégezni. Az összehasonlítás valós anyagon történt: Hollandia és Izrael közös vízhasznosítási programja keretében megjelent egy cikk az *Agricultural Water Management c.* folyóirat 1982. évi 5. számában, szerzői: Th. M. Boers és J. Ben-Asher, címe: *A review of rainwater harvesting* (Az esővízkitermelés áttekintése). A cikk 105 bibliográfiai hivatkozást tartalmaz, ezeket a szerzők hagyományos "kézi" kereséssel találták.

Egy későbbi időpontban gépi keresést végeztek ugyanerre a témára, megállapítandó, hogy melyek azok az irodalmi források, amelyeket a kézi módszerrel nem találtak meg. A gépi keresést úgy végezték el, hogy az egyes felhasznált adatbázisok tartalma és a keresés költségei összevethetőek, értékelhetőek legyenek.

#### A vizsgált adatbázisok

Hogy a lehető legteljesebb eredményre jussanak, minden olyan adatbázist bevontak a keresésbe,

amelyről gyanítható volt, hogy releváns anyagot tartalmaz, nevezetesen:

- ◆ WATER RESOURCES ABSTRACTS (WRA)
- ◆ AGRICOLA
- ◆ COMMONWEALTH AGRICULTURAL BUREAUX ABSTRACTS (CAB)
- ◆ AGRIS
- ◆ PASCAL (francia)
- ◆ AQUALINE
- ◆ BIOSIS
- ◆ COMPENDEX
- ◆ GEOARCHIVE
- ◆ GEOREF
- ◆ LIFE SCIENCE COLLECTION
- ◆ NTIS
- ◆ COMPREHENSIVE DISSERTATION INDEX

A felsorolt 13 adatbázisban szabad keresőszavakból álló, eltérő keresőprofilokkal haladtak végig, a keresőszavak és kombinációik megválasztása az adatbázis tárgykörétől, a változtatás az első keresés eredményétől függött. A kísérlet logikáját a következőképpen alakították ki:

1. *A keresőszavak I. halmaza* a négy első egymással logikai VAGY, ezek a két utolsóval ÉS kapcsolatban:

WATER	(víz)
RAINWATER	(esővíz)
PRECIPITATION	(csapadék)