

## Szakértői rendszerek a tudományos és műszaki információ terjesztésében

### Tudományos és műszaki információs szakértői rendszerek

Az utóbbi években számos szakértői rendszert (más néven döntéselőkészítő, döntéstámogató, konzultációs rendszert, angolul: expert system) fejlesztettek ki. Néhány tipikus alkalmazás: DENDRAL — molekulaszintézis, szerves kémia; MYCIN — orvosi konzultáció; PROSPECTOR — ásványi lelőhelyek feltárása, geológiai; CRIB — számítógépes (hardver) hibák felderítése és javítása; PLCO — ürrepülés.

A szakértői rendszereket olyan kommunikációs eszköznek tekinthetjük valamely terület szakembereinek egy csoportja számára, amely optimalizálja az információ áramlását és növeli a csoport szaktudását. Mint az elnevezésből is következik, a szakértői rendszerek segíteni tudnak problémák megoldásában

- ◆ megoldási lehetőségek kikeresésével vagy generálásával,
- ◆ az adott körülmények között legjobb megoldás kiválasztásával,
- ◆ egy-egy konkrét helyzetben többféle eljárás felkínálásával.

Mindez szoros kapcsolatban van a tudományos és műszaki információ terjesztésének, cseréjének és hasznosításának folyamataival a kutatásban és a termelésben egyaránt. A tudományos és műszaki információval kapcsolatos döntési problémák főleg a következő területeken lépnek fel:

- ◆ tudományos kutatások tervezése, kezdeményezése, kibontakoztatása, kísérletek megtervezése, az eredmények értelmezése stb. terén,
- ◆ a kutatási és az ipari központok közötti információcsere, különösen gyártási és technológiai problémák megoldása terén,
- ◆ a modernizálás, az új gyártmányok tervezése terén, szabadalmi információk hasznosítása és szabadalmak elbírálása során, és így tovább.

Ezeknek a területeknek a döntési problémái adják a *tudományos-műszaki információs szakértői rendszerek* potenciális szakterületeit. A jelenlegi számítógépes tudományos-műszaki információs rendszerek a felhasználó számára releváns dokumentumok, illetve gyakrabban az ilyen dokumentumokra való hivatkozások visszakeresésére alkalmasak, a relevancia megállapítása igen merev kritériumok alapján és elég nyers módszerekkel történik. A jövő információs szakértői

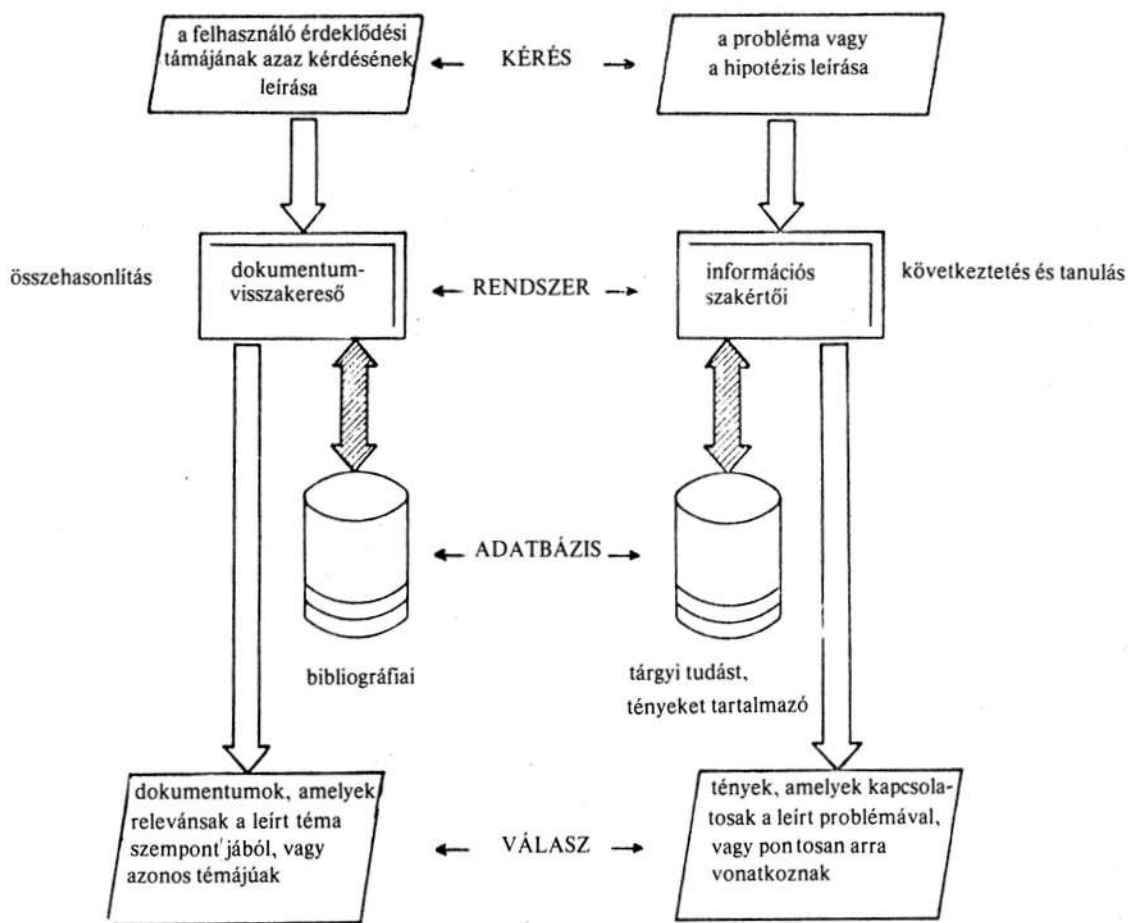
rendszere *igazi információkereső (nem dokumentumkereső) rendszer* lehet, amely — az 1. ábra szerint — a felhasználói kérdéseket problémaleírásnak vagy hipotézisleírásnak tekinti, egy tárgyi ismereteket, tényeket tartalmazó tudásbázis tartalmával összeveti, majd következtetési folyamatok kiindulópontjává teszi. A felhasználó válaszképpen a problémával kapcsolatos, arra vonatkozó tényeket kap.

Az információs szakértői rendszerek jellegzetes feladatai értelmezés jellegűek és diagnózis jellegűek lesznek. Értelmezésen rendszerint adatok elemzését értjük a jelentésük meghatározása céljából, a diagnózis pedig hibakeresés (élő vagy élettelen) rendszerekben, potenciálisan zajos adatok értelmezése alapján. Az információs szakértői rendszerben az adatelemzés beérheti a meglehetősen pontossággal, a teljességet azonban szigorúan biztosítani kell. A pontatlanságon ugyanis felhasználói visszacsatolással sokat lehet javítani, de a teljesség hiányát gyakorlatilag lehetetlen kompenzálni.

Mint a legtöbb szakértői rendszerben, az információs szakértői rendszerben is három szintet különböztetünk meg: az adatok, a tudásbázis s a vezérlés szintjét. Az *adatszint* tartalma az éppen megoldás alatt lévő problémákra és a megoldási folyamat éppen aktuális szakaszára vonatkozó tudás. A *tudásbázis* a rendszer szakterületének speciális tudásanyagát tartalmazza. Ezt használja fel a rendszer, amikor egy-egy problémán 'gondolkodik'. Rendszerint operátorok vagy eljárások formájában van megadva. (Az operátorok és eljárások alkalmazása megváltoztathatja az adatokat, és így finomíthatja a rendszer tudását.) A *vezérlés* szintje rendszerint program, amelyik döntéseket hoz, hogy mikor és hogyan kell használni a specifikus problémamegoldó tudást, például mely operátorokat vagy eljárásokat és hogyan kell alkalmazni.

### Információk (tárgyi ismeretek, tények) a tudományos és műszaki publikációkban

A tudományos és műszaki publikációk nyelvének a köznyelvtől eltérő stílusa és szintakszisa van. A sajátosságok folytán lehet azon gondolkodni, hogy az információs szakértői rendszerekben a tudományos és műszaki nyelvet használjuk — természetes kódként — a tudás reprezentálá-



1. ábra Dokumentum-visszakereső vs. információs szakértői rendszer

sára. A természetes nyelv grammatikájának és szemantikájának formalizálása terén az utóbbi években elért haladás is támogatja ezt a lehetőséget.

A szakértői rendszerek adatbázisai számára a legjobb információforrások az elméleti, módszertani vagy kísérleti témájú tudományos dolgozatok és az új tervek, új gyártmányokat vagy új technológiákat bemutató műszaki közlemények. Az adatelőkészítés fő feladata az adott írás lényegének azonosítása és kivonatolása, vagyis a szöveg azon állításainak és részleteinek kiválogatása, amelyek az eredmények pontos meghatározását és leírását tartalmazzák. A legtöbb esetben elég a kivonatot, a bevezetést és a végkövetkeztetéseket elemezni.

A jó kivonatnak a lehető legrövidebb formában tartalmaznia kell a dolgozatban közölt összes tényt és konklúziót, valamint fel kell hívnia a fi-

gyelmet az új adatokra. Az adott szakterület általánosan elfogadott terminológiáját kell használnia. Az új fogalmakat és szimbólumokat röviden meg kell magyaráznia. A tárgyra vonatkozó leglényegesebb kiegészítő információkat is tartalmaznia kell, és le kell írnia például az új elméletek, folyamatok vagy technikák jellemzőit, a kísérleti módszertan alapelveit, az újonnan kapott numerikus értékeket, stb.

Jó kivonat mellett a bevezetés és a végkövetkeztetések csak segítségül szolgálnak a jobb megértéshez. A bevezetés értékes információforrás lehet még arra vonatkozóan, hogy milyen viszony van a dolgozatban közölt és az adatbázisban már tárolt tények között.

A kiválogatott szövegdarabokat nyelvi szempontból még ellenőrizni kell, esetleges többletműségeiket meg kell szüntetni, megfelelő formába kell önteni őket az adatbázisba való bevétel előtt. Végeredményben az egy bizonyos tényre

vonatkozó információt természetes nyelvű közlésekkel reprezentálhatjuk, egy fő közlés az adott tény megjelenését hangsúlyozza, egy vagy több járulékos közlés pedig leírja a tényt.

*Ezt az eljárást is az indexelés egy fajtájának lehet tekinteni, amelyet azonban – szemben a dokumentumok indexelésével – az információ indexelésének nevezhetünk.*

Megfontolható persze a közlemények teljes szövegének kivonatolása. Az eljárás bonyolultabb, mert az információ mintegy széjjel van szórva az egyes részekben, és ami még fontosabb: olyan tárgyi ismeretek, ill. tények kiemelésére kerülhet sor, amelyek fontosak az indexelő, de nem a szerző szempontjából. A jó kivonatok minden olyan adatot tartalmaznak, amelyekre szükség van egy információs szakértői rendszerben.

### Az információs szakértői rendszer adatbázisa

Ahol a tárgyi ismereteket, tényeket természetes nyelvű mondatokkal reprezentáljuk, ott nem lehet nagy mértékben formalizálni. Azok a legalmasabb struktúrák, amelyek a legjobban őrzik az információ szemantikai jegyeit. Jelenleg ennek a követelménynek leginkább a szemantikai hálók felelnek meg. A szemantikai háló csomópontjai szemantikailag független információ-fragmentumok, szavak vagy kifejezések. Más megközelítésben, a csomópontokban objektumokra (pl. könyv, János), a tényekre (Brazília egy ország) vagy eseményekre (János találkozott Máriával) vonatkozó tudás összegeződik. A csomópontok más csomópontok osztályait is reprezentálhatják (pl. objektumok és tények vagy események osztályait). Minden csomópontnak neve (címkéje) van, a csomópont típusától függően a név vonatkozhat egyedi elemekre vagy az elemeknek egy osztályára. A csomópont jelentése nincs explicit módon megadva, hanem a csomópontok közötti relációkban fejeződik ki. Ezek a relációk a valós világban előforduló függőségeket tükrözik. Azok az elemek, amelyek egy bizonyos csomóponttal közvetlen kapcsolatban vannak, a csomópont szűk környezetét alkotják, a szűk környezet csomópontjaival közvetlen kapcsolatban lévő további csomópontok pedig a tág környezetét.

Az információs szakértői rendszer szemantikai hálója többek között tény-, kifejezés-, szó- és osztálycsomópontokat tartalmazhat. A ténycsomópontok reprezentálhatják a fő és járulékos közléseket (l. fentebb). Ezek a csomópontok a sze-

mantikai háló legalacsonyabb szintjét alkotják. A kifejezéscsomópontok a közlésekben előforduló kifejezéseket reprezentálják. A hierarchia legmagasabb szintjét alkotó szócsomópontok a természetes nyelv szavainak felelnek meg. Az osztálycsomópontok azonos típusú elemeket egyesítenek, és a szavak grammatikai vagy szemantikai jellemzőit jelenítik meg (igék osztálya, előljárók osztálya stb., vagy szinonímák osztálya, azonos tövű szavak osztálya stb.), továbbá jellemezhetik ténycsomópontok egy-egy csoportját is például aszerint, hogy a létezésnek a tényét, vagy a megjelenésnek a tényét stb. tükrözik.

A csomópontokat címkézett, irányított élek kötik össze, amelyek grammatikai és szemantikai viszonyt reprezentálhatnak, ill. egy csomópontot hozzárendelhetnek a csomópontok egy meghatározott osztályához.

### Az információs szakértői rendszer működése

A felhasználó kérésében azaz a probléma-specifikációban szereplő fogalmakat össze kell hasonlítani a háló elemeivel, az összehasonlításnak általában ki kell terjednie a háló minden szintjére és a csomópontok minden típusára. A hálón belüli mozgás stratégiáját a működés folyamán, az éppen összehasonlított fogalmak (szűk és tág) környezetének elemzése alapján lehet meghatározni. Adott fogalom környezetének elemzése során, a szituációtól függően, vagy a specifikusabb, vagy az általánosabb, vagy a szemantikus egyenértékű fogalmakat vesszük figyelembe. A szemantikai hálóban mindig a specifikusabb fogalmak felé kell haladni, a legrosszabb esetben az egyenértékűek felé. Az általánosabb fogalmak felé csak akkor haladjunk, ha a felhasználó részéről visszacsatolás van, tehát ő döntheti el, meddig lehet és hasznos általánosítani. Csak akkor lépünk egy általánosabb fogalom felé, ha más lépésre nincs lehetőség, vagyis ha a specifikusabb vagy egyenértékű fogalmak környezetének nincs közös része.

A folyamatnak akkor van vége, amikor kiválgattuk a hálóból a kérésnek megfelelő összes közlést. Ezeket vagy megkapja a felhasználó, vagy inputadatként kezeljük őket a probléma további vizsgálatához. Más esetekben átmehetnek egy további ellenőrzésen és/vagy levezetésen, és ennek az eredménye lesz a probléma megoldása. A megoldás a szituációtól függően lehet egyszerű igen/nem válasz vagy hosszú következtetés, ill. döntés is.

A szakértői rendszerek növekvő térhódítása a tudományos-műszaki információ terjesztésében negatív mellékhatásokat is eredményezhet, így várhatóan egyre több bonyolult jogi problémával kell szembenézni.

### Automatikus keresésközvetítők (intelligens interfészek) az online rendszerekben

Az információkereső rendszerek teljesítményét nemcsak a rekordok tartalmának és/vagy a keresési módszereknek a változtatásával lehet növelni, hanem a kereső és a rendszer közé iktatott automatizmussal is, amely az ember és a rendszer közötti párbeszéd minőségét javítja. A két módszer természetesen együtt is alkalmazható. Ez a dolgozat a második lehetőséggel foglalkozik.

Az online információszolgáltatók, így a Dialog vagy Data-Star abban érdekeltek anyagilag, hogy nőjön a keresések száma, ezért közvetlen hozzáférést biztosítanak a felhasználóknak a releváns adatbázisokhoz. Így kívül reked a közvetítő szakember, aki tudja, hol és hogyan kell keresni, noha a tevékenységére továbbra is szükség van. Vajon elláthatja-e a közvetítő funkcióit egy automatikus keresésközvetítő rendszer, az ún. intelligens interfész?

#### Az interaktív információkeresés fő lépései

1. A potenciális kérdezőben tudatosul egy HIÁNY, egy információ iránti IGÉNY.
2. Aki ténylegesen kérdezővé válik, az IGÉNY-t kérdés formájában fejezi ki.
3. A kérdező a kezdeti kérdések tisztázására törekszik, akár egyedül, akár közvetítő segítségével dolgozik.
4. A tisztázott kérdést le kell fordítani, azaz a megfelelő terminológia használatával és a megfelelő keresési stratégia alkalmazásával olyan kifejezést kell csinálni belőle, amelyet az információkereső rendszer kezelni tud.
5. Sor kerül az első keresésre, megszületik az eredmény.
6. A kérdező (és a közvetítő) megvizsgálja az eredményt, ennek alapján a kérdést módosítja. A 4.-6. lépések ismétlődnek, amíg már nem kell módosítani a kérdést.
7. A kiértékelés összeveti a keresés eredményét és a kezdeti IGÉNY-t.

/NOWAK, E. J. — SZABŁOWSZKY, B. F.:  
*Expert systems in scientific information exchange = Journal of Information Science*, 8. köt. 3. sz. 1984. p. 103—111./

(Szöllősy Éva)

#### A felhasználó igényei

Üzemszerű, kompetens és emberszabású automatikus keresésközvetítő rendszer létrehozása még sok munkát követel. A jelenlegi rendszerek, amelyek egyáltalán dialógust tudnak folytatni az emberrel, általában csak egyszerű kérdésekre adnak pontos választ, és az egyszerű párbeszéd is csak egy merev grammatikához ragaszkodva lehetségesek. A rendszerek tudnak egyszerű felvilágosítást kérni, de ha a felhasználó válasza nem világos, azt megérteni nem képesek.

Ha két beszélgető ember nem érti meg egymást, a társalgás nem jut csödbe: rövid tisztázó szub-dialógus után a beszélgetés folytatható ott, ahol abbamaradt. Ilyen tisztázó dialógus kibontakoztatására a gépi rendszerek képtelenek. Ehhez előbb tudnunk kellene, hogyan vannak szervezve a felnőtt embernek a világról való ismeretei, és hogy ez a tudás hogyan kerül felhasználásra.

A mesterséges intelligencia kutatása már sok energiát áldozott az emberrel kommunikálni képes intelligens interfészek létrehozására. Hayes és Reddy [13] kellemes társalgó rendszereket sürgetnek, ez a kellemesség (graceful interaction) sok különböző képességből áll össze:

- ◆ *rugalmas nyelvelemzés* — a rendszer legyen képes kezelni az idiómákat, a nyelvtani hibákat, a töredékes közléseket, tehát mindazt, ami a természetes nyelvben előfordulhat,
- ◆ *szilárd kommunikáció* (robust communication) — olyan stratégiák kellenek, amelyek biztosítják, hogy a hallgató fogadja és helyesen értelmezze a beszélő kijelentéseit,
- ◆ *összpontosító mechanizmus* — a rendszer legyen képes követni a beszélgetés menetét, és számon tartani, miről van, ill. volt szó,
- ◆ *magyarázó mechanizmus* — a rendszer legyen képes megmagyarázni, mit tud és mit nem tud csinálni, mit csinált, mint próbál csinálni és miért, mégpedig direkt kérdésre válaszolva is és a kommunikáció pillanatnyi csődje esetén is,