

A felmérés módszere és eredményei

Az NFAIS-hoz tartozó 15 R és I szolgáltatás képviselőivel telefonos interjúkat készítettek. Közülük 8 a természet- és műszaki tudományok területén, 4 a társadalomtudományokban, 3 pedig a humán tudományokban fejti ki tevékenységét. Három szolgáltatás 1981-ben 100 ezernél több dokumentumot dolgozott fel, nyolc 10 ezer és 100 ezer között, négy pedig 10 ezernél kevesebbet. Hárman már legalább öt éve nyújtanak online szolgáltatást is, hárman öt évnél kevesebb ideje, négynek nincs nyilvánosan, online módon hozzáférhető adatbázisa.

Valamennyi válaszból az *előfizetés változásának hasonlósága* tűnik szembe. A változások trendjeire az adatbázis-előállítók részéről való reagálás azonban erősen eltérő.

Mindegyik válaszolónak az az érdeke, hogy termékeik előállításában és terjesztésében korszerűbb módszereket vezessenek be. Valamennyinek van már, illetve tervezéssel géppel olvasható adatbázis; egyesek mikroformás publikációt terveznek. Némely szolgáltatás videotex, videolemez vagy optikai lemez használatát helyezi kilátásba.

A már évek óta nyilvános online szolgáltatás keretében elérhető adatbázisok előállítói a használat gyors felfutását jelentették. Egyikük sem érte el a csúcspontot, tehát a használat további növekedését várják.

A nyomtatott termék előfizetésében 1976–1977-től kezdve átlagosan évi 2–5%-os lemorzsolódásról számolnak be. Az egyik adatbázis-előállító egy gyorsan növekvő szakirodalommal és komoly ipari háttérrel rendelkező, „divatos” szakterületet fed le, ahol inkább az előfizetés növekedése, semmint csökkenése lett volna várható. Nőtt viszont a teljes adatbázisokból készített, egyes résztémákat lefedő, kisebb kiadványok előfizetése, ezeket ugyanis egyének is vásárolják saját zsebből.

Az előfizetési díjak fokozatosan növekednek, hogy kompenzálják az inflációt és a lemorzsolódást. A bevétel tehát állandó vagy kissé nő. Az R és I szolgáltatók bevételeinek nagy része ma is a nyomtatott termékek eladásából származik, ez a trend azonban változhat. A géppel olvasható adatbázisok eladásából, bérbeadásából és jogdíjaiból származó bevételek százalékosan is gyorsabban nőnek, mint a kiadványok bevételei. Egyes szolgáltatók az „elektronikus termékek” árát úgy emelték, hogy a kétféle termék tekintetében paritást érjenek el.

A nyomtatott termékekre előfizetők csökkenésének nem kizárólagos oka az online adatbázisok elérhetősége. Ennek egy bizonyítéka az, hogy az online hozzáférést még nem, vagy öt évnél rövidebb ideje nyújtó szolgáltatók ugyanolyan mértékű előfizetés-csökkenésről számolnak be, mint a már régebben online elérést is lehetővé tevő adatbázis-előállító intézmények. A legnagyobb csökkenés annál a szolgáltatónál volt, amelynek nincs online elérhető adatbázisa. A lemorzsolódás okai között

a gazdasági nehézségek, a mikroformák és a reprográfia terjedése, a felhasználók igényeinek változása stb. szerepelnek.

Az R és I szolgáltatások (adatbázisok) előállítóinak válasza a kiadvány előfizetőinek lemorzsolódására *nyugyon eltérő*. Némelyikük csak a nyomtatott termékben közöl kivonatot, az online szolgáltatásban nem, mások csak a kiadvány előfizetőinek engedik meg az online hozzáférést, megint mások az árakkal operálnak. A leginkább hangoztatott válasz: fejlesszünk ki új termékeket és szerezzünk új piacokat. Egyes adatbázis-előállítók sikeresen önköltségüket csökkenteni. Nagy részük választotta a szolgáltatás tökéletesítésének útját, pl. új adatelemek (adatmezők) bevezetését, a termékskála bővítését, az elosztás könnyítését.

A felmérésből kiderült, hogy a nyomtatott termékek és az online elérhető adatbázisok kölcsönhatása nem mindig negatív vagy gazdasági természetű. Az online hozzáférés új felhasználó-csoportokat hozott az irodalmazás számára. Számos gazdasági és ipari szakember szívesebben fordul a géppel olvasható adatbázisokhoz, mint nyomtatott változatához. A tudományos kutatók is sok helyütt szívesebben veszik igénybe a multiadatbázis-keresés nyújtotta lehetőségeket az interdiszciplináris irodalmazásban.

A kétféle termékcsoport versengése előtérbe hozta a felhasználók érdekeit és igényeit. A megkérdezettek közül az utóbbi években négyen végeztek felmérést adatbázisuk felhasználói körében. Mások is egyre inkább élnek a felhasználói visszacsatolás lehetőségeivel.

A nyomtatott és az elektronikus termék önálló, egymástól független fejlesztése is a trendek közé tartozik. Ezzel jár, hogy az egyik közegen elérhető olyan termék, amelynek nincs megfelelője a másik közegben.

A címben feltett kérdésre tehát a következő választ lehet adni: a kezdeti, látszólag negatív hatást követően a kétféle termék kedvezően befolyásolja egymást, versengésük az információszolgáltatás korszerűsítését váltja ki.

/SPERR, I. L.: Online searching and the print product: impact or interaction? = Online Review, 7. kötet. 5. sz. 1983. p. 413–420./

(Roboz Péter)

Az online keresés pszichológiája

Az online számítógépes információkeresés az „új kommunikációs forradalom” legfőbb vívmánya, hiszen az információrobbanás hatásaival az automatizált adatfeldolgozás és a telekommunikáció legfejlettebb eszközeit összekapcsolva veszi fel a harcot. Amíg azonban a terület megdöbbentően gyors fejlődését és társadalmi-gazdasági kihatásait számbavéve egyenesen „online jelen-

ségről” beszélhetünk, egy sor probléma felderítetlen, megoldatlan, elsősorban az ember–számítógép kapcsolat vonatkozásában. Az online keresési folyamatra irányuló számos vizsgálat közös tanulsága például az, hogy a *használó–rendszer interakció alacsony színvonalú*: a keresők egyszerű stratégiákkal dolgoznak, és nem használják ki a technológia biztosította visszacsatolási, iteratív, heurisztikus problémamegoldási lehetőségeket; elhanyagolt a keresési stratégiák és a kereséskor lezajló pszichés folyamatok kutatása.

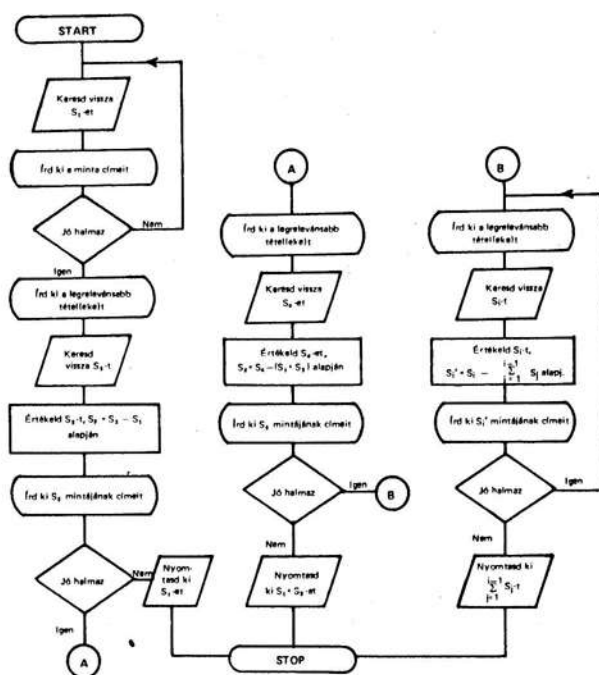
Interakció, redundáns hivatkozások, információ túlterhelés

Az online kereséskor folytatott ember–gép párbeszéd egyik nehézsége a *visszakeresett halmazok értékelésében és a legjobb halmazok végső kimenet (output) céljára történő kiválasztásában rejlik*. A kereső ugyanis általában úgy jár el, hogy visszahív egy, mondjuk 100 tételes halmazt. Az ebből vett minta alapján eldönti, hogy ez releváns-e számára. A címek azonban a kérdés új módjára (például szinonimákra) hívhatják fel a figyelmet, így a módosított keresőképpel megismétli a lekérdezést, és tegyük fel, ismét egy 100 tételes halmazhoz jut. Ha a minták alapján mindkét halmazt relevánsnak találja, a két halmazt egy VAGY operátorral összeköti és kinyomtatja. Azt azonban nem tudhatja – az emberi memória és adatfeldolgozási készség korlátai miatt –, hogy a második halmaz valamennyi releváns tétele szerepel az elsőben, s így (80%-os pontosságot feltételezve) az új halmazzal húsz irreleváns tételt „nyert”. Igaz, hogy releváns tétel nem veszett el, a keresés minősége mégis romlott: a pontosság csökkent, és a teljesség változatlan maradt. Az irreleváns tételek szaporodása (különösen ha két halmaznál többet kapcsolnak össze a VAGY-gyal) olyan információ túlterheléshez vezet, amely a felhasználót a keresés befejezésére vagy leegyszerűsítésére készíti.

Mindez elkerülhető az ún. *zárt ciklusú relevancia klaszteráló algoritmus* segítségével (1. ábra), amely a Boole-i NEM operátorral megoldja azt, hogy a korábbi jó (elég pontos) halmaz(ok)éval egybeeső tételek a később lekérdezett halmazban már a képernyőn való megjelenítéskor se szerepeljenek. Így biztosítható, hogy az új halmazok értékelése *egyedi és új információtartalmuk alapján történjen*, és kiessenek a redundáns (egyszer már visszakeresett) hivatkozások.

Az online keresés hatékonyságának és átlátásának elősegítése a negáció felhasználásával

A zárt ciklusú relevancia klaszteráló algoritmus hatását két öt fős, állatorvostanhallgatókból álló csoport keresésén mérték le. A kísérleti és a kontroll csoport



1. ábra Egy zárt ciklusú relevancia klaszteráló algoritmus

A blokkdiagram a halmazoknak a redundáns hivatkozások kiiktatásával történő összevetésére mutat egy algoritmizált módszert. S_1 ténylegesen nem az első visszakeresett halmaz, hanem az első olyan logikai formula, amely a kereső pontossági kritériumának megfelelő, kinyomtatandó, tehát jó halmazt eredményez

kétórás előadás és 90 perces gyakorlat keretében, valamint egy, a *DIALOG Users Manual* (használói kézikönyv) 1. és 3. fejezetét, *Bates* keresési taktikáit és hat mintakeresést tartalmazó kézikönyv segítségével sajátíthatta el az online keresés fogásait a DIALOG rendszerben, az egyik mezőgazdasági témájú adatbázisban. Az egyetlen különbség az volt, hogy a kísérleti csoport tantervében a klaszteráló algoritmus is szerepelt.

Bár a külső tényezők kiszűrésére nem volt mód, az eredmények azt sugallják, hogy *az algoritmus ismerete és megértése előnyös*: a három kérdésre végzett, egyenként harminc perces keresés során a kísérleti csoport mind a keresés hatékonyságában, mind az eljárás- és utasításrendszer, illetve a keresési stratégia (logika) hibátlanágában messze felülmúlta a kontroll csoportot.

Az ember–számítógép interface: analóg/bináris kommunikációs szakadék

A számítógépes irodalomkutatás alapvető problémája, hogy két olyan rendszernek (a szubjektív, analóg embernek és az objektív, bináris számítógépnek) kell párbeszédet folytatnia, amelyek *homlokegyenest ellenkező adat-*

feldolgozási módot követnek. A számítógép ugyanis – képtelen lévén a kérdés értékelésére, a kérdező szándékának felfogására – pontosan arra válaszol, amit kérdeztek tőle, s ez nem esik feltétlenül egybe azzal, amit a kérdező óhajtott.

A problémát Zadeh [1] az *életlen halmazok* bevezetésével közelítette meg. Az életlen halmaz jelen esetben egy olyan halmazt jelent, amely (a halmazhoz tartozás pontatlan meghatározása következtében) irreleváns elemeket (hivatkozásokat) tartalmaz. A kereső számára ez azzal a következménnyel jár, hogy a kívánt teljességet és pontosságot csak ismételt lekérdezésekkel érheti el. Mivel a kezdő keresőkép általában pontatlan és nem elég átfogó, a visszacsatolást kihasználva a kereső módosított kérdésekkel fog próbálkozni, a lehetséges és optimális kombinációk kialakítására a kombinatorikát és a faktoranalízist felhasználva. Meg kell küzdenie azonban a teljesség és a pontosság inverz kapcsolatával is.

Szekvenciális és egyidejű kognitív folyamatok

A teljesség és a pontosság együttes optimalizálása csak a fentiekben említett párbeszéd segítségével érhető el. Ebben az emberi információfeldolgozás két módja (a lineáris és időbeli szekvenciális, illetve a térbeli és párhuzamos egyidejű, szimultán feldolgozás) közül a szimultán játszik nagyobb szerepet, mert ez szükséges a visszahívott halmazok összehasonlításához, a keresés eredményeinek egységesítéséhez, a kombinatorikai sorozatok faktorelemzéséhez. Bármilyen jó szimultán képességgel rendelkezen is a kereső, az információ túlterhelés csak technológiai segítséggel, a redundancia automatikus csökkentésével kerülhető el.

Miller [2] mutatott rá arra, hogy az emberi elme igényli az információk térbeli orientációját. A térbeli képességek viszont a matematikaiakkal függenek össze. Így, mivel az online kereséskor használt absztrakt rendszerekből hiányzik a térbeli összetevő, és ezt a keresőnek kell kialakítania, arra lehet következtetni, hogy az interaktív keresés *könnyebb azok számára, akik fejlett matematikai készségekkel rendelkeznek.* Ezek hiányában a kereső olyan keresési rutinokra fog támaszkodni, amelyek nem alkalmazkodnak rugalmasan az adott folyamathoz, problémához, és nem lesznek képesek a keresőkérdés hatékony átalakítására, illetve a halmazok manipulálására (összevetésére, rangsorolására stb.).

A negáció

A Boole-i NEM operátor használata a számítógépes keresés kulcskérdése. A keresők azonban többnyire kerülnek alkalmazását, és ennek ismét csak lélektani okai vannak. Az online rendszerek használói kézikönyvei

szükszavúan és általában csak a fogalomkizárás (például: gyümölcs, de NEM alma) eszközeként említik, megjegyezve, hogy ha mind a kívánt, mind a kizárt témával foglalkozik a dokumentum, legyen bár releváns, a NEM alkalmazása miatt elvész. Ez megerősíti a tagadás amúgy is negatív konnotációját. A NEM használata ellen hat továbbá az a Watson [3] által bizonyított pszichológiai tény is, hogy a logikai műveletek közül a negációt a legnehezebb felfogni, és a tagadó mondatok több gondolkodási időt és intellektuális erőfeszítést igényelnek, mint az állítások.

Ezek a tényezők eredményezik azt, hogy a negációt a redundáns hivatkozások kiküszöbölésére sem használják fel. Az ismétlődő hivatkozásokat tartalmazó halmazok azonban – mint erre korábban utaltunk – zavart keltenek, ezért a kereső vagy lemond az iterációról az első jó halmazt elfogadva, vagy egyszerűen VAGY-gyal kapcsolja össze a jó halmazokat, ahelyett, hogy megkülönböztetésükkel és módosításukkal a végső kimenet pontosságát javítaná.

Az ember kereső magatartása: képek, tervek helyett

A *Keresési és megoldási tervek* című tanulmány [4] rámutat, hogy a legtöbb ember a „van” és a „kellene” közötti szakadékra összpontosít, azaz a problémahelyzetet egy statikus képben rögzíti a maga számára, ahelyett, hogy a keresés vagy a problémamegoldás alternatív lehetőségeit járná végig, szisztematikus terveket kidolgozva. Ez nyilvánvalóan érvényes azokra is, akik online keresést végeznek, és ebből következik, hogy a használók képzésében nem elég a keresés alapvető lehetőségeit (a visszacsatolást, a Boole-logikát stb.) ismertetni, hanem be kell mutatni azokat a keresési stratégiákat is, amelyek szisztematikus és tervszerűen alkalmazzák ezeket az eljárásokat, és így csökkentik az emberben keletkező kognitív feszültségeket.

Következtetések

Az online keresésben rejlő lehetőségek teljes kihasználását csak a hatékony ember-gép párbeszéd teszi lehetővé. Az online rendszerek biztosítják az azonnali visszacsatolást, így a heurisztikus és építőjáték-szerű előrehaladást, ami – ha a használó a megfelelő technika birtokában van – sikert, és az online kereséssel kapcsolatos kedvező beállítódást eredményez. A visszacsatolás ugyanakkor a keresés sikerének a záloga is, hiszen a kérdező mindig tisztában van a lekérdezett halmazok viszonylagos értékével, és tudja, hogy a halmazok bővítésekor milyen, esetleg kapcsolódó tételek nem jelennek meg az outputban. Az ismertetett algoritmus elsősorban összetett, iteratív keresések rendszerezésére szolgál, de

javíthatja a kérdező Boole-logikai készségeit, és az online keresési folyamat megértését, áttekintését is.

Hivatkozások

1. ZADEH, L. A.: Fuzzy sets = Information and Control, 8. köt. 1965. p. 338–353.
2. MILLER, G. A.: Psychology and information = American Documentation, 19. köt. 3. sz. 1968. p. 286–289.
3. WATSON, P. C.–JOHNSON-LAIRD, P. N.: Psychology of reasoning. London, B. T. Batsford, Ltd. 1972.
4. MILLER, G. A.–GALANTER, E.–PRIBRAM, K. H.: Plans and the structure of behavior. New York, Henry Holt, 1960. p. 159–176.

/VIGIL, P. J.: The psychology of online searching = Journal of the American Society for Information Science, 34. köt. 4. sz. 1983. p. 281–287./

(Orbán Éva)

Az online szolgáltatások helyzete Olaszországban

Olaszországban a 70-es években kezdett kialakulni az online kultúra, elsősorban a nemzetközi szervezetek révén. Ezek közül elsőként kell említeni az ENSZ szakosított szervezeteit, köztük az Élelmiszer és Mezőgazdasági Szervezetet (FAO), amelynek Rómában van a központja, valamint a bécsi központú Nemzetközi Atomenergia Ügynökséget (IAEA). Az előbbinek az AGRIS (Mezőgazdasági Információs Rendszer), az utóbbinak az INIS (Nemzetközi Nukleáris Információs Rendszer) adatbázisa építésében, majd hasznosításában Olaszország is közreműködött. Sokáig az egyetlen nemzetközi adatbázisközpont, amelyhez az olaszországi felhasználók is hozzáférhettek, az Európai Űrkutatási Szervezet Információvisszakereső Rendszere, a jól ismert ESA-IRS volt, amelynek ESANET hálózatán keresztül lehetett online kapcsolódni a Róma melletti Frascatiban működő számítógéphez és a rendszer mintegy 40 adatbázisához.

Az Euronet/DIANE használata

Az olasz posta- és távközlésügyi minisztérium felelős a hálózat használatáért az országban, hasonlóan az Európai Gazdasági Közösség többi tagállamához. Az Euronet római koncentrátora és milánói csomópontja a minisztérium Euronet-központjának a hatáskörébe tartozik mind az üzemeltetés, mind pedig a használat tekintetében, de itt végzik a statisztikai felméréseket is.

Az Euronet/DIANE rendszer hasznosításának elősegítését a Nemzeti Kutatási Tanács (CNR) keretében szervezték meg. Létrehozták a DIANE olaszországi forrástájékoztató központját, amelynek feladatai – többek között – a felhasználók és közvetítők oktatása, a rendszerre vonatkozó információk gyűjtése és terjesztése, az információk kérések továbbítása a megfelelő szolgáltató központokhoz, valamint tanácsadás. A központ 16 európai host-számítógéppel áll kapcsolatban, és hozzá tud férni az amerikai Dialog és Orbit rendszer adatbázisaihoz is, a TYMNET hálózaton keresztül.

Az Euronet hálózat olaszországi hostjai

Host alatt az olyan számítógépet értjük, amely egy vagy több adatbázist vagy adatbankot tárol és kezel, és a hálózathoz kapcsolódva, annak bármely felhasználója rendelkezésére áll. Ilyen host az olasz Legfelsőbb Bíróság adatbázisát szolgáltató *Italgire-Find* rendszer, ill. az azt kezelő Univac 1100/81 típusú számítógép. A különféle jogi, igazságügyi, törvényhozatali stb. információkat feldolgozó rendszerhez országszerte 200 terminállal kapcsolódhatnak, intézményekből, kutatóintézetekből, egyetemekről, tudományos társulatoktól.

Az Italgire-Find rendszerhez 20 online hozzáférhető adatbázis tartozik, ezek négy csoportra bonthatók:

- törvényhozási adatbázisok (az olasz törvények és rendeletek teljes szövege 1972-ig, illetve címe 1860-ig visszamenőleg);
- az olasz Legfelsőbb Bíróság polgári és büntető eljárásainak anyaga, különleges esetei;
- bizonyos különleges bíróságok ügyei és határozatai;
- jogi, igazságügyi és törvényhozási szakirodalmi adatbázisok.

Egy következő fontos host a CNUCE, a pisai egyetem Országos Egyetemi Elektronikus Számítástechnikai Központja*, amelyet 1974-ben a CNR vett át. Eredményei közé tartozik egy öt csomópontos számítógéphálózat, valamint két adatbáziskezelő rendszer, amely az Euronet/DIANE rendelkezésére áll, továbbá egyéb szoftverszolgáltatásai. A CNUCE különféle adatbázisokat is hozzáférhetővé tesz az Euronet/DIANE számára, amelyek a következő csoportokba sorolhatók:

- négy törvényhozatali adatbázis;
- európai nemzetközi (politikai, gazdasági és információs) kapcsolatokra vonatkozó öt adatbázis;
- két művészeti adatbázis.

* A CNUCE központot és bizonyos szolgáltatásait a budapesti SZTAKI is hasznosítja online hozzáféréssel. – *A szerk.*