



"Átlátszó" információkereső rendszerek

Az átlátszó (transparent) információkereső rendszerek fogalma azt kívánja kifejezni, hogy a felhasználó a keresési tevékenység bonyolult láncolatán, mint az üvegen átlát, mivel közvetlenül a keresési eredmény tárul a szeme elé. A felhasználó felteszi a kérdést a keresőrendszernek és visszakeresi mindazt az adatot vagy információt, amelyek kérdésére a választ tartalmazzák, vagy amelyekben a válasz alapul. Közben nem érzékeli azokat a bonyolult folyamatokat, amelyek a kérdés feltevése és a végeredmény szolgáltatása között a rendszerben végbemennek. A világ összes ismeretének átlátszó rendszere még messze van, azonban számos részletet már megoldották vagy fáradoznak a megoldáson. Ezeket a rész megoldásokat a kapuszolgálatok (gale-way), az előfeldolgozó (front-end) rendszerek, az illesztések (interface) és a közvetítő rendszerek fogalmával fejezik ki.

Az alábbiakban az átlátszó információkereső rendszerek szükségességéről és az erre irányuló kutatásokról lesz szó.

Az adatbázisok és az online rendszerek fejlődése

Az utóbbi időben hallatlanul megnőtt az adatbázisok és az online szolgáltatók száma. Az 1976-ban számon tartott 310 adatbázis helyett 1985-ben több mint 3000 szöveges, numerikus és grafikus adatbázist ismertünk, a géppel kereshető rekordok száma ez idő alatt 50 millióról 1680 millióra nőtt. 1976-ban főleg bibliográfiai adatbázisok léteztek, ma az adatbázisok 36%-a numerikus vagy grafikus jellegű. A szöveges adatbázisok közül pedig a bibliográfiaiak mellett megtaláljuk a teljes szövegű, az útmutató (directory) és forrástájékoztató (referral), valamint az enciklopédikus jellegű adatbázisokat is. Az adatbázis-készítők száma világszerte mintegy 1200, az online szolgáltatók száma pedig meghaladja a 600-at.

Az átlátszó rendszerek létjogosultsága

Az adatbázisok, a keresőrendszerek és a keresési módszerek széles skálájában való eligazodás megkönnyítése sokak szerint az egységesítés, a szabványosítás nyújtana. A szabványok bevezetésére azonban éveket kell várni, és a rendszerek, az adatbázisok rendkívüli változatossága nem is mindig engedi meg a szabványosítást. A szabványok használata önkéntes, sokszor jár azzal, hogy jól begyakorolt eljárásokat, sok évi gyakorlatot kell felrúgni és újakkal felváltani. Ehelyett a felhasználóknak, akik sok adatbázisban, különböző rendszerekben és más-más módszerekkel kénytelenek kereséseiket végezni, egy *látszólag szabványos* megoldási biztosítanak a front-end, az interfész, a közvetítő rendszerek és a kapuszolgálatok. Mindezek egyenként is, együttesen is az online keresés eltéréseit, sokrétűségét láthatatlanná, vagyis a keresőrendszert átlátszóvá lehelik a felhasználók számára.

Számos kutatási program foglalkozón és foglalkozik az állászatosság számítástechnikai és távközlési megoldásaival. Ezek közé tartozik az egységes hozzáférés problémája több rendszerhez, az adatbázisok automatikus kiválasztása és szolgáltatóinak meghatározása, a keresés automatikus segítése, az output-formátumok átalakítása, a duplikátumok kiszűrése a keresés eredményéből, az automatikus bekapcsolódás és bejelentkezés, valamint a többi ehhez hasonló program. Az ilyen problémakörökben elért eredményekre használják a különböző szerzők a már idézett négy szakkifejezést, vagy éppen az átlátszó rendszer kifejezést. Valamennyinek közös a célja: az online keresés bonyolultságai és sokféleségét csökkenteni. Ilyen értelemben találónak tartjuk az *átlátszó keresőrendszer* elnevezést, valamennyi fenti funkció közös terminológiájaként.

Az átlátszó rendszert lehetővé tevő számítástechnikai – többnyire szoftver – termékek között találunk mikroszámítógépi programcsomagokat helyi

használatra, front-end rendszereket (hardver és szoftver) a szolgáltatóközpont használatára és olyan rendszereket, amelyek a felhasználóitól, illetve a szolgáltatótól független helyen működnek. Egyes rendszerek előfeldolgozási végeznek, mások a keresés folyamata közben teljesítenek feldolgozási feladatokat, megint mások utófeldolgozást látnak el

Az online keresési segítő automatikus funkciók

Az online keresés menete több résztevékenységre bontható fel:

- a keresőkérdés megvitatása;
- az adatbázis(ok) kiválasztása;
- a szolgáltató és az adatátviteli hálózat meghatározása a fentivel összhangban;
- a kulcsszavak, az információkereső nyelvi elemek (deszkriptorok, kötött tárgyszavak, osztályozási jelzetek stb.) megkeresése (A magyar keresőknek további gondot okoz a megfelelő idegen nyelvű szakterminológia megtalálása. — A ref);
- a keresési stratégia felépítése;
- bekapcsolódás a szolgáltatóhoz, majd a megfelelő adatbázisokhoz;
- a stratégia bevitele;
- a közbenső találatok értékelése;
- a végső találati halmaz online vagy offline ki nyomtatása, esetleg letöltése a saját számítógépre;
- az eredmény értékelése;
- a megfelelő dokumentumok megrendelése vagy - ha lehet - online hozzáférése.

Ugyanezt a folyamatot esetleg több rendszeren kell megismételni. Ha numerikus adatokra, idősorokra vagy modellező rendszerekre van szükség, még tovább bonyolódik a helyzet.

Ezek a tevékenységek mind meghatározón funkciókra bonthatók és a legtöbb funkcióban megvan az automatizálás lehetősége (egyeseké már meg is valósult). Ha ez sikerül, bármely automatizált funkció vagy ezek kombinációja közelebb visz a végső célhoz: átlátszó információkereső rendszerekhez. A funkciókat a legegyszerűbben a keresési megelőző, a keresés közbeni és a keresési követő funkciókként csoportosíthatjuk. Az automatizálás szempontjából azonban más vonatkozásban kategorizáljuk őket.

Az automatizálható funkciók kategóriái

Az információkereséssel kapcsolatos elemi funkciók a következő fő kategóriákba sorolhatók:

1. automatikus átalakítók.
2. automatikus kapcsolatirányítók,
3. automatikus választók,
4. automatikus értékelők.

Az *automatikus átalakítók* körébe a következő funkciók tartoznak:

- Elérési protokollok: A rendszer elérése megvalósítható egy *B* rendszeren keresztül és fordítva. Ide tartozik az elérési protokoll nyelvének átalakítása egyikről a másikra.
- Parancsnyelvek: A rendszer parancsnyelvével egy *B* rendszeren is kereshetünk és fordítva. Hasonlóképpen a parancsnyelv természetes nyelve is konvertálható (angol, francia stb. nyelvű parancsok átalakítása egymásba).
- Rendszerüzenetek átalakítása.
- Adatbázisok információkereső nyelve: az 1. adatbázis kötött keresőszavai átalakíthatók a 2., 3. stb. adatbázis keresőszavaivá.
- Egy adatbázis szabad keresőszavai átalakíthatók az adatbázis kötött információkereső nyelvére és fordítva.
- Az adatbázis szabad keresőszavai az egyik természetes nyelvről másik nyelv keresőszavaivá alakíthatók át.
- Rövidítések, kódok, szimbólumok, szinonimák stb. kölcsönös átalakítása.
- Audiobemenet átalakítása digitális adatokká és viszont (éppen úgy, mint a beszéd felismerésben, illetve hangszintézisben).

Az *automatikus kapcsolatirányítók* közé tartoznak az automatikus bekapcsolódási és bejelentkezést (log-on, log-in) és egy adatbázisba való bekapcsolódási végző programcsomagok. Ugyancsak ide tartozik a kereső személy automatikus irányítása olyan útmutatóhoz (az útmutató-adatbázisok katalógusából), amely a keresőkérdésnek, a kért adatoknak leginkább megfelel.

Az *automatikus választók* a következő funkciókat képesek ellátni:

- A keresőkérdésnek megfelelő adatbázisosztály kijelölése (ebben számításba jönnek a kereső személy jellegzetességeiből kiinduló tanuló rendszerek).
- A megfelelő adatbázis kiválasztása, főleg a keresőszavak gyakorisága, szinonimái stb. alapján.
- Az online szolgáltató kijelölése a szükséges adatbázis(ok), a használati díjak, a visszamenőlegesség, a felhasználó jártassága, a kisegítő szolgáltatások stb. alapján.
- Az adatátviteli hálózat kiválasztása.

- A keresőkérdésben használ! keresőszavak kiválasztása olyan átlátszósági segédletekkel, amelyek automatikusan meghatározzák a szinonim és a kapcsolódó kifejezéseket. Az ilyen célra készült programok természetes nyelvű szavakat, kötött keresőszavakat, szinonimákat, hierarchikus kapcsolatokat, csonkolásokat, szógyakorosságait stb. képesek kezelni.
- Alkalmazói programcsomagok az adatok egyesítésére, manipulálására, statisztikai elemzésekre, modellezésre stb.
- Az adatok optimális formátummeghatározása.

Az *automatikus értékelők* sok funkciói képesek ellátni.

- A visszakeresésre szám információs egység színijének illesztése a kérdést fellelvők igényeihez, a rájuk vonatkozó tárolt adatok alapján.
- Az információs egységek minősítése, az értékes és kevésbé értékes egységek elkülönítése.
- A keresés automatikus finomítása asszociatív módszerekkel.
- Az automatikus visszakeresési funkciók telepítési helye változó lehel. A rendszer szervezői határozzák meg, hogy melyik funkciót hová építsék be. Az automatikus visszakeresést segítő átlátszósági eszközökéi lehel telepíteni:
- felhasználóoldali intelligens terminálon, mikro-, mini- vagy közepes számítógépen;
- a felhasználói és a szolgáltatótól egyaránt független adatátviteli vagy kapcsolódó számítógépen, erre a célra beállított berendezéseken;
- szolgáltatóoldali számítógépen;
- a felhasználótól és szolgáltatótól független al-rendszereken, több helyen levő elosztott adatokkal;
- a fentiek kombinációján.

Az automatikus keresést segítő állászsósági rendszerek architektúrája attól függ, hogy egy vagy több funkciót bíztak-e rá. Egy tipikus átlátszósági eszköz, a kapuszolgálat például automatikus bejelentkezést, adatbázis-választást, parancsnyelv-átalakítási, üzemformátum-átalakítást és más funkciókat tehet lehetővé, hogy az egyik szolgáltató rendszeréi egy másikéhoz legye hasonlóa. A kapuszolgálat architektúrája ezek szerint változik. Az adminisztratív és irányítási funkciók is – az adatokhoz hasonlóan – azonos vagy különböző helyeken lehelnek telepítve.

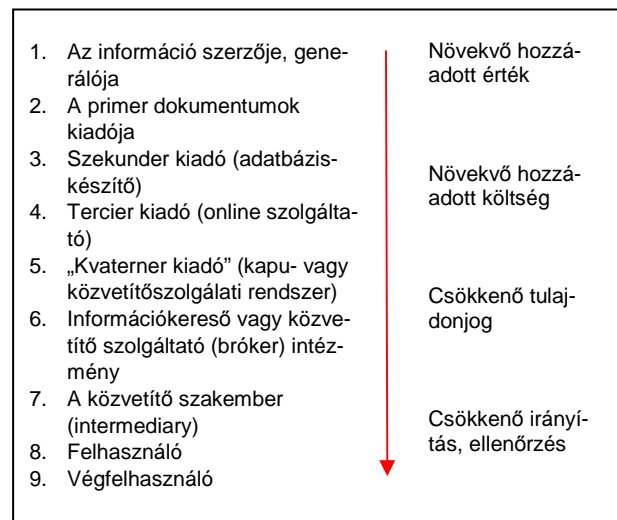
Problémák, feladatok

A hajlékonylemezes és kompaktlemezes adatbázisok vagy állományok egyrésztől, a kapuszolgálá-

tok megléte másrésztől felveti a centralizáció vagy decentralizáció kérdéseit. A terminálok növekvő intelligenciája, a mikroszámítógépes adat manipuláció terjedése a helyi feldolgozás – decentralizáció – irányába hat, de az intelligencia beépítése egy kapuszolgálati számítógépbe a centralizálási segíti. Utóbbi egyszersmind átvesz a szolgáltatótól addig nála centralizált funkciókat, így a kapuszolgálat egyes irányítási feladatok ellátását is átveszi.

Az információ magánjellege, a feltett kérdés bizalmas kezelése minden felhasználó jogos igénye. A magánemberek vagy a vállalatok nem vennék jó néven, ha bárki megtudná, hogy milyen kérdéseket tesznek fel. Az adatbázis-készítők, a szolgáltatók, a kapuszolgálatok és az adatátviteli hálózatok ezt nem tekintik problémának, mert nem áll érdekükben a kérdések „lehallgatása”, és erről biztosítják a felhasználókat is. Más kérdés, hogy mindnyájan érdekeltek abban, hogy megismerjék a felhasználói piac érdeklődését, vagyis hogy melyik felhasználói réteg vagy csoport milyen típusú információt keres. Ennek oka egyszerűen a termékfejlesztés, a marketing.

Az *adatbázis-használati lánc* minden eleme, amely az információhoz értéket ad hozzá, érdekelt a piac azonosításában. A lánc a következő:



Ezek közül a kvaterner kiadóval foglalkozunk, mint az átlátszóság egyik legfontosabb elemével. Ebben foglaltuk össze valamennyi típusú kapuszolgálatot vagy közvetítő rendszert – az utóbbit megkülönböztetve a keresési közvetítéstől –, amely valamely felhasználó (a 7., 8. vagy 9. közül) és a szolgáltató között áll és biztosítja köztük az elektronikus kapcsolatot. Ez képviseli azokat a fenti

funkciókat, amelyek megkönnyítik vagy integrálják a felhasználók hozzáférései több online rendszerhez, egységes, állatszó (láthatatlan) módon. Nyilvánvaló, hogy a láncban lefelé haladva nő az információhoz hozzáadott érték és költség, csökken viszont az információs egységek kézben tarthatósága, kontrollja.

Vertikális kapuszolgálatok azok, amelyek egy diszciplínában, egy szakterületen nyújtanak integrált hozzáférést, szemben a horizontális kapuszolgálatokkal, amely két vagy több szolgáltatót kapcsol össze. A vertikális kapuszolgálat teszi lehetővé a felhasználóbarát, átlatszó elő- és utófeldolgozó online keresőszolgáltatást mezőgazdasági, jogi, orvosi stb. témában, legyen az bibliográfiai, numerikus vagy bármely típusú adatbázis és szolgáltatása bármelyik szolgáltatóközpont. Az ilyen vertikális kapuszolgálatot a szolgáltatóközpontok vagy adatbázis-készítők csoportjai hozhatják létre.

A CD-ROM technika és a kapuszolgálatok kapcsolata

Az újonnan bevezetett technikai eredmények közül egyesek az online szolgáltatókat fenyegetik. A kapuszolgálati rendszerek is hasonló fenyegetésnek vannak kitéve, mert az online szolgáltatóktól függenek.

A CD-ROM (kompaktlemez hordozón tárolt és forgalmazott adatbázisok – Compact Disc Read-Only Memory) technika alkalmazása arra vezethet, hogy az adatbázis visszakerül a felhasználó szervezet kezébe, mint annak idején a mágnesszalagos SDI-szolgáltatások esetében. Viszont a szolgáltatóközpontban végzett online keresés megszabadítja a felhasználói szervezeteit egy se reg gondtól. Ezért nem valószínű, hogy ismét lemondanak a központi erőforrások megosztott használatának és a sok adatbázishoz egy helyen való hozzáféréseinek előnyeiről. Azzal a lehetőséggel, hogy nagy adatbázisokat lehel kompaktlemezekre megkapni, azokkal házon belül manipulálni és azokból tetszés szerint visszakeresni, csak olyan szervezetek fognak élni, amelyekben kevés adatbázisra, de azokból nagyszámú keresésre van szükség. A sok adatbázist használók vagy a kis volumenű használók körében a CD-ROM-hordozós adatbázisok nem fognak elterjedni.

Ma még nem ismeri a CD-ROM-hordozó hatása az online szolgáltatókra. Ha a népszerű, nagy adatbázisok esetén a nagy fogyasztó felhasználók CD-ROM keresésre térnek át, akkor a szolgáltatók nem tudnak kellő bevételre szert tenni ahhoz, hogy a kisebb használatú adatbázisokat a jelenlegi áron fenntartsák. Ekkor az várható, hogy a kis adatbázisok használata tovább csökken, elsősorban a kevésbé pénzes felhasználók részéről. Az eredmény ezek szolgáltatásának megszűnése lehel, ami viszont azokat sújtja, akik az ilyen adatbázisokat értékesnek tartják.

A CD-ROM-hordozó fő előnye az archiválás lehetősége, különösen a belső gyűjtemények és teljes szövegű dokumentumok újszerű kezelése. A gyakori aktualizálást igénylő állományok és a tranzakciós adatbázisok nem alkalmasak CD-ROM-hordozó használatára.

Következtetések

Az átlatszó információkereső rendszerek iránti igényi már az online szolgáltatások kezdetén felismerték. A rájuk fordított kutatást az USA-ban az állam finanszírozza. Az ötletek egy részét profitorientált szolgáltató vállalatok meg is valósították, ezáltal segítenék a felhasználókat és növelték az adatbázisok használatát az online rendszereken. Az *átlatszóságot segítő* módszereket még egyik nyilvános online szolgáltató sem valósította meg teljesen. Nyilván a még be nem vezetett átlatszósági eszközöket a legnehezebb megvalósítani. A kutatók ma is intenzíven foglalkoznak például a keresést segítő szakértői rendszerek kidolgozásával, amelyek mintegy utánozzák a szakemberek gondolkodását. A nehezebb funkciók ellátásához – pl. a természetes nyelvű szöveges adatbázisok torzításmentes lefordítása más nyelvre vagy nyelvekre – a mesterséges intelligencia módszerei szükségesek, amelyek még váratnak magukra.

Becslésünk szerint az átlatszó információkereső rendszerek bemutatón funkcióinak 85%-át 1990-re üzembe helyezik.

WILLIAMS, M. E.: Transparent information systems through gateways, front ends, intermediaries, and interfaces. = Journal of the American Society for Information Science, 37. köt. 4. sz. 1986. p. 204-214./

(Roboz Péter)