

önálló vállalkozói, vagy független kutatói minőségükben eljáró keresők szereznek a rendszer alkalmazásából).

### Szakértő rendszerek

A valós környezetben végzendő kutatások szükségességét növeli az a körülmény, hogy teauruszok és más hasonló keresési segédletek valószínűleg hamarosan láthatatlanná válnak az olyan szakértő rendszereken belül, amelyek lehetővé teszik a keresőkérdések természetes nyelven történő megfogalmazását. E szempontból fontos az ilyen rendszerek készítőinek megállapítása [7]: a teaurusz rendkívül fontos eszköz annak biztosítására, hogy a rendszer sok különböző módon megfogalmazódott azonos kérdés nyomán ugyanazt a tartalmat keresse.

A teaurusznak a szakértő rendszerekben való alkalmazása külön kérdés, egyesek integrálni akarják az ismeretbázisba, mások inkább megőriznék önálló segédeszköznek.

\* \* \*

A keresés hatékonyságát annyi külső változó befolyásolja, hogy valószínűleg nincsenek általános érvényű optimális megoldások. Még az sem állítható, hogy mindig a két módszer együttes alkalmazását kell ajánlani. Leginkább arra van szükség, hogy a módszerek működését egyre jobban megértsük, mégpedig pontosan leírt környezetekben vizsgálva őket, amelyekben a releváns változót pontosan azonosítani lehet.

### Irodalom

- [1] HENZLER, R. G.: Free or controlled vocabularies. = International Classification, 5. köt. 1978. p. 21– 26.
- [2] HERSEY, D. F. et al.: Free text word retrieval and scientist indexing and retrieval. = Journal of Documentation, 27. köt. 1971. p. 167– 183.
- [3] WEINBERG, H. B.: Multiple sets of human indexing for civil engineering documents: Comparison of structure and occurrence rates in full text. = Science and Technology Libraries, 2. köt. 3. sz. 1982. p. 13– 33.

- [4] DUCKITT, P.: The value of controlled indexing systems in online full text databases. = Proceedings of the 5th International Online Information Meeting, Oxford, Learned Information, 1981. p. 447– 453.
- [5] MARKEY, K. et al.: An analysis of controlled vocabulary and free text search statements in online searches. = Online Review, 4. köt. 1980. p. 225– 236.
- [6] JOHNSTON, S. M.: Effect of thesaurus indexing in retrieval from machine-readable databases. = Quarterly Bulletin of the International Association of Agricultural Librarians and Documentalists, 27. köt. 1982. p. 90– 96.
- [7] FENICHEL, C. H.: An examination of the relationship between searching behavior and searcher background. = Online Review, 4. köt. 4. sz. 1980. p. 341– 347.
- [8] TATALIAS, J.: Attitudes and expectations of potential and user online searchers. = Proceedings of the National Online Meeting, New York, 1985. p. 457– 462.
- [9] TENOPIR, C.: Full text database retrieval performance. = Online Review, 9. köt. 2. sz. 1985. p. 149– 164.
- [10] BRYANT, J. H.– TERAPANE, J. F.: Online searching in the US Patent and Trademark Office. = World Patent Information, 7. köt. 1– 2. sz. 1985. p. 133– 138.
- [11] ROTHMAN, J.: Is indexing obsolete? = Feinberg, H. (ed.) Indexing Specialized Formats and Subjects. Metushen, New Jersey and London, Scarecrow Press, 1983. p. 22– 23.
- [12] DUBOIS, C. P. R.: The use of thesauri in online retrieval. = Journal of Information Science, 8. köt. 2. sz. 1984. p. 63– 66.
- [13] PEREZ, E.: Text enhancement, controlled vocabulary versus free text. = Special Libraries, 73. köt. 3. sz. 1982. p. 183– 192.
- [14] GASTALDY, S. B.: Les thesaurus de recherche: des outils pour l'interrogation en vocabulaire libre. = Argus, 13. köt. 2. sz. 1984. p. 51– 58.
- [15] FEINBERG, H.: The thesaurus in indexing and searching – I. 11., p. 260– 281.
- [16] WILLETTS, M.: Investigation of the relation between terms in thesauri. = Journal of Documentation, 31. köt. 3. sz. 1975. p. 158– 184.
- [17] BERNSTEIN, L. M.– WILLIAMSON, R. E.: Testing of a natural language retrieval system for a full text knowledge base. = Journal of the American Society for Information Science, 35. köt. 4. sz. 1984. p. 235– 247.

/DUBOIS, C. P. R.: Free text vs. controlled vocabulary; a reassessment. = Online Review, 11. köt. 4. sz. 1987. p. 243– 253./

(Szöllősy Éva)

## A hasonlósági keresés úttörője: a duplumrekordok kizárása a Dialognál

Az online információkeresés mostanáig mindig pontos egyezéssel keresést jelentett, vagyis a keresőkifejezés és a kikeresett dokumentumban előforduló kifejezés – a csonkolás vagy maszkolás megengedte szabadsággal – pontosan meg kellett egyezzen egymással. Az ennél általánosabb hasonlósági keresésre, vagyis arra, amikor a kikeresett

dokumentumban a keresőkifejezéshez hasonló, de attól azért többé kevésbé eltérő kifejezés előfordulását várjuk el, már voltak kísérletek. Az itt ismertetethez hasonló duplumkizárással például főleg mikroszámítógépes útfeldolgozás keretében próbálkoztak. Az itt ismertetett eljárás azonban az első, amely a hasonlósági keresést nyilvános keresőrendszerben valósítja meg. – A ref.

A Dialog online információkereső rendszer *OneSearch* lehetősége módot ad rá, hogy a kívánt információkeresést egyidejűleg több adatbázisban végezzük el. Ez égetően veti fel, hogy az egyazon dokumentumra vonatkozó több adatrekordot azonosítani és kezelni tudjuk. A hasonló tematikájú adatbázisok között ugyanis nagy tartalmi átfedés várható,

vagyis gyakori, hogy az együtt keresett adatbázisok közül több is tartalmazza egy bizonyos dokumentum adatrekordját. Erről az átfedésről fogalmat alkothattunk az 1. táblázatból, amely úgy készült, hogy egyazon keresést két adatbázisban is elvégeztünk. Olyan témát választottunk, amely kis találatsszámot ad. A kapott találatokat azután manuálisan hasonlítottuk össze\*.

1. táblázat

## Tartalmi átfedés különböző adatbázisok között

A adatbázis	B adatbázis	A B adatbázissal közös adatrekordok aránya Az A adatbázisban (%)	Az A adatbázissal közös adatrekordok aránya a B adatbázisban (%)
Agricola	Agris	10,9	18,1
Agricola	Biosis	22,7	8,1
Agricola	CAB	20,3	35,9
Biosis	CAB	4,8	23,9
Biosis	FSTA	2,2	20,2
Biosis	SciSearch	18,3	32,9
Compendex	Inspecc	20,4	19,9
Embase	Medline	70,3	52,8
Embase	SciSearch	55,4	32,6
Medline	SciSearch	28,2	14,8

Az egy dokumentumra vonatkozó több adatrekord gyakran lényegében azonos információt tartalmaz. Ugyanakkor viszont eltéréseket találhatunk a címben és az egyéb bibliográfiai adatokban, eltérőek lehetnek a kulcsszavak és a tartalmi kivonat.

Az információkeresés eredményét általában úgy kell átválogatnunk, hogy egy dokumentumra vonatkozólag csak egy találatunk maradjon. Ezt manuálisan elvégezni nagyon időigényes. Ha viszont automatizálni akarjuk a válogatást, akkor ehhez két kérdésre kell előzetesen választ találnunk:

- ▶ Mit tekintünk duplumrekordnak?
- ▶ Hogyan azonosíthatjuk a duplumrekordokat?

**A duplumrekordok azonosítása**

Bibliográfiai adatbázisokban a duplumrekordok azonosítása azt jelenti, hogy feltárjuk a *hasonlóságot* egyazon adat különböző ábrázolásai között.

Egy adott szinten például a következő problémákkal találkozhatunk magunkat szemben:

- ▶ Azonosnak tekintünk-e két bibliográfiai tételt, ha azonos a szerzőjük és a címük, de eltérő a forrásuk? (Például az egyik folyóiratban jelent meg, a másik konferenciakiadványban.)
- ▶ Azonos-e a két tétel, ha megegyezik a szerzőjük és a címük, de eltérések vannak olyan forrásadatokban, mint a megjelenés éve vagy az oldalszámok? Alacsonyabb szinten például a következő problémák merülhetnek fel:
- ▶ A két adatbázisban eltérő a szerzőnevek ábrázolásmódja. Hogyan dönthetjük el, azonos-e a két szerző?

- ▶ A két adatbázisban eltérőek a cím ábrázolási szokásai. (Például a kis egész számokat az egyik betűvel írja, a másik számjeggyel; eltérően alakítják a karakterkészlettel ábrázolható formára az alsó és felső indexeket, a különleges karaktereket; eltérően fordítják le a más nyelven megjelent cikk címét.) Hogyan dönthetjük el, azonos-e a két cím?

Az ilyen problémák állnak a középpontban, ha jól működő azonosítási algoritmust akarunk kidolgozni.

**Kulcsgenerálási megoldások**

Bibliográfiai adatok esetén a duplumrekord-detekálás eddig megvalósított megoldásainak középpontjában a kulcsgenerálás áll. Olyan kulcsokat kíséreltek meg generálni, amelyek egyértelműen azonosítják az adatokat, ugyanakkor viszont kellően érzéketlenek a tipográfiai eltérésekre. A várakozás az, hogy az ilyen kulcsok szerinti rendezés feltárja a megegyező kulcsokat, ezen keresztül a megegyező dokumentumokra vonatkozó bibliográfiai rekordokat.

\* Az átfedés két hasonló témakörű adatbázis között erősen *témafüggő*. A mezőgazdasági tematikájú CAB és a biológiai tudománnyal foglalkozó BIOSIS adatbázis között például nagy átfedés várható a termesztett növények mutációival kapcsolatban, de alig várható átfedés, ha a téma a szántás technológiája, illetve a kísérleti patkányok tartási körülményei. A SciSearch adatbázis olyan témaköröket is tartalmaz (pl. csillagászat), amelyek egyáltalán nem fordulhatnak elő az orvosi vagy biológiai adatbázisokban. A kis mintán talált átfedés tehát a szerző állításával ellentétben nem tekinthető jellemzőnek az adatbázisok egésze közötti átfedésre, mindössze jellemző példának arra, hogy egyes témákban *lehet* nagy az átfedés. – A ref.

A kulcsgenerálás első lépése a kiindulásul szolgáló adatmezők kiválasztása. Több e tárgyban végzett vizsgálat mutatja, hogy megfelelő feldolgozás esetén jó jelölt a kulcsgenerálás kiindulópontjául a cím és az alcím. Kiegészítő információként szolgálhat a publikálás éve, a szerző és az oldalszámok (legalábbis a kezdő oldalé) [1, 2]. Ha rendelkezésünkre áll, akkor megbízható adat a CODEN is. Az ISBN és a Library of Congress katalógusszám azonban elfogadhatatlanul magas hibaszázalékot adnak [3].

Ha a kiinduló adatmezőket már kiválasztottuk, akkor az ezekből történő kulcsgenerálásának két módja használatos:

- ▶ A leképezéses módszerek előnye a gazdaságos tárkihasználás. Megfelelő leképező függvényekkel ugyanis nagyon kompakt kulcsokat hozhatunk létre [4, 5].
- ▶ A másik lehetséges megoldás a karakterszelekciós módszer. A kulcsba kerülő karakterek kiválasztása rendszerint pozíciójuk alapján történik. A belőlük összeállított karakterfüzér lesz a kulcs. A legjobb kiválasztási eljárás meghatározása entrópiamegfontolások alapján történhet [2].

Online alkalmazásokra a karakterszelekciós módszer tűnik előnyösebbnek, mert

- ▶ egyszerű;
- ▶ olyan kulcsok generálhatók vele, amelyek bizonyos mértékig őrzik a rendezési sorrendet, ezzel ellenőrzési lehetőséget adnak;
- ▶ a gyakori hibák elemzésére alapozva viszonylag érzéketlenné tehető a tipográfiai hibákra.

A karakterszelekciós eljárások általában a lehető legkevesebb karakterre korlátozódnak, és rendszerint bizonyos szavak első karaktereit választják ki, bár egyes megoldások ettől eltérnek.

A karakterkiválasztást általában a kiválasztás hatékonyságát javító manipulációk előzik meg. A legkézenfekvőbb előkészítő műveletek a következők lehetnek:

- ▶ a kisbetűk konvertálása nagybetűvé;
- ▶ a központozás, a kötőjelek és a felesleges többletszóközök eltávolítása;
- ▶ a vezető névelők eltávolítása.

Ezek az előkészítő lépések általánosak a karakterszelekciós eljárásokban. Egyes javasolt eljárások további előkészítő lépéseket is alkalmaznak:

- ▶ Valamennyi névelő, előljárószo és kötőszó eltávolítása.
- ▶ Bizonyos gyakori szavak (pl. *ARE, IS, BE*) eltávolítása.
- ▶ A belső magánhangzók eltávolítása, a kettős más-salhangzók közül a második eltávolítása.
- ▶ A kulcsgenerálásban felhasznált karakterkészlet csökkentése mondjuk a tizenhat legkevesebbé gyakori karakterre.

Néhány ismert karakterkiválasztási módszert foglal össze a 2. táblázat.

## 2. táblázat

### Karakterszelekciós módszerek

Forrás	Kiválasztási eljárás
OCLC	A cím első szavának első három betűje, második és harmadik szavának első két betűje, negyedik szavának első betűje
Slach [6]	A publikálás éve, az első szerző vezetéknevének első négy betűje, a kezdő oldalszám
Bowker	Az első szerző vezetéknevének első négy betűje, a cím első négy betűje
MacLaury [2]	A cím első szavának első és harmadik betűje, második és harmadik szavának első betűje, hátulról harmadik szavának első betűje, utolsó előtti szavának első betűje, utolsó szavának negyedik és utolsó előtti betűje
Brooks [1]	A publikálás éve, a kezdő oldalszám, az első szerző vezetéknevének első betűje és első hat mássalhangzója, első keresztnevének kezdőbetűje, a folyóirat neve első négy szavának első két-két betűje, a cím első és utolsó négy-négy mássalhangzója

### Műveletek a kulcsokkal

A kulcsok alkalmazása közvetlen és közvetett módszerekre osztható.

A *közvetlen* módszer során a rekordokat kulcsaik szerint rendezzük, az egy dokumentumra vonatkozó rekordok kulcsai így egymás mellé kerülnek. Ennek a módszernek az eredményessége nagyon érzékenyen függ a mezőkiválasztás és kulcsgenerálás választott módszerétől.

A *közvetett* módszer során a kulcsok szerinti rendezés arra szolgál, hogy a rekordokat "valószínű megfelelők" szerinti csoportokba soroljuk. Az így kiválasztott csoportokon belül azután részletesebb vizsgálatot végzünk.

A Brooks [1] által javasolt eljárásban a kezdeti rendezés a publikálás éve és a kezdő oldalszám szerint történik. Ezután további vizsgálatnak az így szomszédossá váló kulcsokat kell alávetni. A további felhasznált mezőket (cím, szerző, folyóirat és kötet) fontossági sorrend szerint súlyozzuk, a súlyok adatbázisfüggőek. A mezők összehasonlítását ezután csökkenő súlyok szerinti sorrendben végezzük, az

összehasonlítás eredményét egy előre eldöntött küszöbérték túlhaladása esetén tekintjük úgy, hogy megegyező rekordokról van szó.

MacLaury [2] szintén kétlépéses eljárást javasol. Az első lépésben a címből és a publikálás dátumából kiválasztott karakterekből képezünk kulcsot, e szerint rendezünk. Azokon a csoportokon belül, amelyekre az így képzett kulcs megegyezik, a cím, a szerző és az oldalszámok alapján döntünk a dokumentumok azonosságáról. A teljes cím szerinti összehasonlításban kapott közelségi értéket, amely itt a második lépésben játszik szerepet, Williams [3] vizsgálja tüzetesebben.

Altom [7] olyan eljárást javasolt, amelyben Brooks [1] módszerét fejleszti tovább. Ebben azonban kiaknázza azt az előnyt, hogy a megcélzott adatbázis (az USA energiaügyi minisztériumának adatbázisa) nagyon szigorúan szabályozott szerkezetű. Ezért a cím, a szerző és az oldalszámok mezőkön kívül felhasználhatja a CODEN, a kutatásjelentésszám, a szabadalomszám, a kiadó, és a folyóiratkötet és -szám mezőket is.

Összefoglalva, a duplummeghatározás legelfogadottabb eljárásai három lépésből állnak:

► Összehasonlítási kulcsok generálása adatcsökkentéssel. Ezek a módszerek olyanok legyenek, hogy ne zavarják meg az elírások és az adatbeviteli előírásokban található eltérések.

► A rekordoknak a valószínűleg egyezők szerinti csoportosítása a generált kulcsok szerinti rendezéssel.

► Az egy csoportba került rekordok részletes vizsgálata a ténylegesen egyezők felderítésére.

### A Dialog módszere

A Dialog rendszer számára kifejlesztés alatt álló duplumdetektálási eljárásban a kulcsokat a cím és a szerző mezőkből képezzük. Az eljárás nem nyújt százszázalékos biztonságot. A hibaarányt kísérletileg ellenőriztük, ennek eredményét a 3. táblázat foglalja össze. A kísérletet úgy végeztük, hogy rokon tematikájú bibliográfiai adatbázisok csoportjait válogattuk össze. Csoportonként húsz-húsz keresést végeztünk. Az így képezett találati halmazokon belül meghatároztuk a duplumrekordokat egyrészt a javasolt gépi eljárással, másrészt manuálisan. A két meghatározás közötti eltérés adja a gépi eljárás hibáját.

Valamennyi adatbáziscsoportra 7% alatt maradt az összes duplumhoz képest a géppel fel nem tártak aránya, és elhanyagolhatóan kicsi, 1% alatti volt a tévesen duplumnak minősített rekordok aránya.

A kísérlet szerint tehát a kidolgozott algoritmus hatékonysága elfogadható. Ez az eljárás elbírja a kisebb eltéréseket a szerzők nevében és a leggyakrabban címszerkesztési szokásokban.

### 3. táblázat

#### Kísérlet a gépi duplumfeltárás ellenőrzésére

Adatbáziscsoport	A találat-halmazok száma	A találat-halmazok átlagos rekord-száma	A talált duplumok száma		Az elszalasztott duplumok		A hamis duplumok	
			manuálisan	géppel	száma	aránya	száma	aránya
Magazine Index Trade and Industry Index Computer Database	20	99	411	405	10	2,43%	4	0,97%
Biosis Agricola CAB Abstracts	20	144	691	658	33	4,78%	0	0%
Computer Database Compendex Inspec SciSearch	20	105	979	933	50	5,11%	4	0,41%
Biosis Medline CA Search	20	75	647	602	45	6,96%	0	0%

### Irodalom

[1] BROOKS, A. A. et al.: A computerized scheme for duplicate checking of bibliographic data bases. = W-7405 -ENG-26 sz. jelentés, Oak ridge National Laboratory, Oak Ridge, Tennessee, USA, 1976.

[2] MACLAURY, K. D.: Automatic merging of monographic data bases. — Use of fixed-length keys derived from title strings. = Journal of Library Automation, 12. köt. 2. sz. 1979. p. 143–155.

[3] WILLIAMS, M. E.: Automatic merging of monographic data bases. — Identification of duplicate records in

- multiple files: the IUCS scheme. = *Journal of Library Automation*, 12. köt. 2. sz. 1979. p. 156–168.
- [4] KNUTH, D. E.: *The art of computer programming*. 3. köt.: *Sorting and searching*. = Addison-Wesley Publishing Co., 1973.
- [5] LYNCH, M. F.: *Variety generation. – A reinterpretation of Shannon's mathematical theory of communication, and its implications for information science*. = *Journal of the American Society for Information Science*, 28. köt. 1977. p. 19–25.
- [6] SLACH, J. E.: *Detection and elimination of duplicates from multidatabase searches*. = *Bulletin of the Medical Library Association*, 73. köt. 3. sz. 1985. p. 234–237.
- [7] ALTOM, D. W. et al.: *Automated duplicate elimination in bibliographic databases*. = Nem publikált dolgozat, Department of Energy, Office of Scientific and Technical Information, Oak Ridge, Tennessee, USA, 1976.
- /BARBARINO, M.: *Similarity detection in online bibliographic databases*. = *Online Information 89. Proceedings (of) 13th International Online Information Meeting, 12–14 December 1989, London, England. Learned Information, Oxford and New Jersey, 1989. p. 111–117.*

(Válasz György)

## Az üzleti adatbázisok fejlődési irányai

Az online információkeresés leggyorsabban fejlődő területe az üzleti adatbázisok köre. Ezt a fejlődést követni szinte lehetetlen.

A következőkben néhány jellemző irányt áttekintve arra teszünk kísérletet, hogy némi képet adjunk arról, hol tartunk és merre haladunk.

### Utazással kapcsolatos adatbázisok

Gyorsan nő az utazással kapcsolatos adatbázisok száma és népszerűsége. Legjellemzőbb példájuk az *OAG Electronic Edition*, amely már több mint tucatnyi szolgáltatóközpontban elérhető. Ez az *Official Airline Guide* számítógépes változata. Puha kötésű nyomtatott változata ott van minden sokat utazó üzletember irodájában vagy aktatáskájában, rendszerint agyonhasználva, mintha a kutya szájából húzták volna ki. Tulajdonosa több mint kétfélmillió légi járat (nemzetközi járatok és az USA belső járatai) közül a nap bármely órájában kiválaszthatja az őt érdeklőket, emellett 29 000 szálloda közül válogathat, le is kötheti a helyét a kiválasztott légi járatra. Nagy népszerűsége folytán ez az adatbázis nyújtja sok üzletember számára az első találkozást az online információkereséssel.

Ugyancsak utazással kapcsolatos adatbázisok a *Mobil Hotel Guide*, a *Visa Advisory Service* és a *Department of State Travel Advisory Service*.

### Teljes szövegű adatbázisok

Fontos fejlődési irány a teljes szövegű adatbázisok előretörése. Egyrészt a szolgáltatóközpontok, pl. a *Dialog* és a *Dow Jones News/Retrieval* egyre több teljes szövegű adatbázist tesznek elérhetővé, másrészt az adatbázisok mind több újság és folyóirat szövegét tartalmazzák.

### Regionális adatbázisok

Az egyesült államokbeli szolgáltatók versenyben állnak a regionális információkat feldolgozó adatbázisok terén. Ilyenek például a *Trade and Industry ASAP*, a *Business Dataline*, ilyen volt a *Predicasts PROMT* adatbázisba rövid önálló lét után beolvadt *Regional Business News*.

Ezek az adatbázisok általában az egyesült államokbeli helyi lapokban és folyóiratokban közölt információkat dolgozzák fel.

### Növekvő adatbázisok

Az online információkeresés felhasználói szeretnek minél kisebb számú adatbázisból minél teljesebb információhoz jutni. Ezért egyes adatbázis-készítők igyekeznek minél átfogóbbá tenni az adatbázisukban feldolgozott információt. Jó példa erre a *Predicasts* nagyon népszerű, számos szolgáltatóközpontban elérhető *PROMT* adatbázisa, amelybe sok olyan információt beolvasztottak, amelyek korábban több külön adatbázisban, egymástól elválasztva voltak csak elérhetőek. Ezt nevezik az adatbázisok előállítói "megafile" koncepciónak.

### Idősor-adatbázisok

Néhány új numerikus adatbázis gazdasági-üzleti adatok idősorait teszi közvetlenül elérhetővé, iparágak (vagy termékek) és földrajzi körzetek szerinti bontásban. Ezen a téren a *Dialog* vállalt úttörő szerepet, amikor a *Chase Econometrics* által készített több mint 10 000 olyan idősort tett hozzáférhetővé a széles közönség számára, amely korábban csak egy szűk felhasználói kör által volt elérhető, közvetlenül az előállítónál.

A cikk megírása után indult és (hazánkban is!) rövid idő alatt nagyon népszerűvé vált idősor-adatbázis a külkereskedelmi forgalmat és árakat a vámnyilvántartások alapján közlő *TRADSTAT*. – A ref.

### Erőteljes fejlesztés a szolgáltatóközpontokban

Az online szolgáltatóközpontok nagyon offenzív szolgáltatási politikát folytatnak az üzleti adatbázisok terén. Ez mennyiségi gyarapítást is jelent (a *Dialog* például háromévenként megkétszerezi a tárolókapacitását, a *Mead*, a *Source* és a *GENIE* gyorsan növelik a náluk elérhető üzleti adatbázisok számát), de fő támadási irányuk a végfelhasználók közvetlen bevonása az online keresésekbe.

A *Dialog* menürendszerű keresőrendszerrel hozott létre az online keresésbe közvetlenül bekapcsolódni