

A CA SEARCH MÁGNESSZALAGOS FELDOLGOZÁS GYORSÍTÁSA

Számítógépes Kémiai Szakirodalomfigyelő Szolgálat a Veszprémi Vegyipari Egyetemen

Egyházy Tiborné—Kertész Miklós

Veszprémi Vegyipari Egyetem Központi Könyvtára—SYSTEM Vállalat

A szolgáltatás alapja 1979. június 30-ig a *Chemical Abstracts Service (CAS)* Ca Con nevű (Chemical Abstracts Condensates) adatbázisa volt. A mágnesszalagos adatbázis a *Chemical Abstracts*-ben ismertetett valamennyi primer dokumentum, mintegy 14 ezer tudományos és műszaki folyóirat válogatott közleményei, könyvek, szabadalmak, kutatási jelentések, doktori disszertációk bibliográfiai adatait és a nyomtatott változat tárgymutatójába bekerülő tárgyszavait tartalmazta. Az 1979. július 1-től forgalomba került CA SEARCH adatbázis a fentiekén kívül még a *General Subject Index* és a *Substance Index* tárgyszavait is magában foglalja, s ezzel a keresőprofilok sokkal árnyaltabb megfogalmazását, a szakirodalom válogatásának finomabb módszerét teszi lehetővé.

A működéshez szükséges PLAN 3 programnyelvű keresőprogramot 1970-ben a KFKI Számítástechnikai Főosztálya dolgozta ki, ICL 1905 típusú számítógépre. A feldolgozás ezen a gépen folyt 1978. december 31-ig, amikor a KFKI ezt a munkát R 40-es számítógépre telepítette át, és kidolgozta az új visszakereső programrendszert, a BINAR-t.

A Veszprémi Vegyipari Egyetem Központi Könyvtára 1979-től részben, 1980-tól pedig teljesen átvitte a feldolgozást a SYSTEM Vállalat R 22-es számítógépére.

Az adatbázisban a keresés szekvenciálisan történik, vagyis újabb dokumentum feldolgozása csak akkor kezdődik el, ha az előző dokumentum már valamennyi profillal összehasonlításra került. A dokumentumnak a profillal való összehasonlítása során a dokumentumnak csak azok az adatai (a dokumentum címe, a szerző neve, a folyóirat címe, a szekciósza-
vakkal, amelyeket a felhasználó kívánt, és a profilszerkesztő a profilban megjelölt.

A feldolgozás gyorsítása a visszakereső programrendszerbe beépített előtét-program segítségével

A megszerkesztett keresőprofilokkal hetente végzünk keresést a CA SEARCH aktuális mágnesszalagjaiból. Az 1. ábra a ráfordítási költségek (a Chemical Abstracts Service-nek fizetendő előfizetési díj, találatsszám és profilszám szerint fizetendő jogdíj, valamint a számítógépes feldolgozás költségei) és a témák számának alakulását mutatja be az évek függvényében.

1979-től kezdve a keresőprofilok számának viszonylagos növekedése mellett *nagyon megnőtt a ráfordítási költségek összege*. Ennek oka a témák számának csekély növekedésén kívül az adatbázis terjedelmének változása, majd pedig a profil-jogdíj bevezetése volt.

Az adatbázis évről-évre növekvő előfizetési díja, a jogdíjak és a hosszabb feldolgozási idők az egy témára eső költségek növekedését eredményezték.

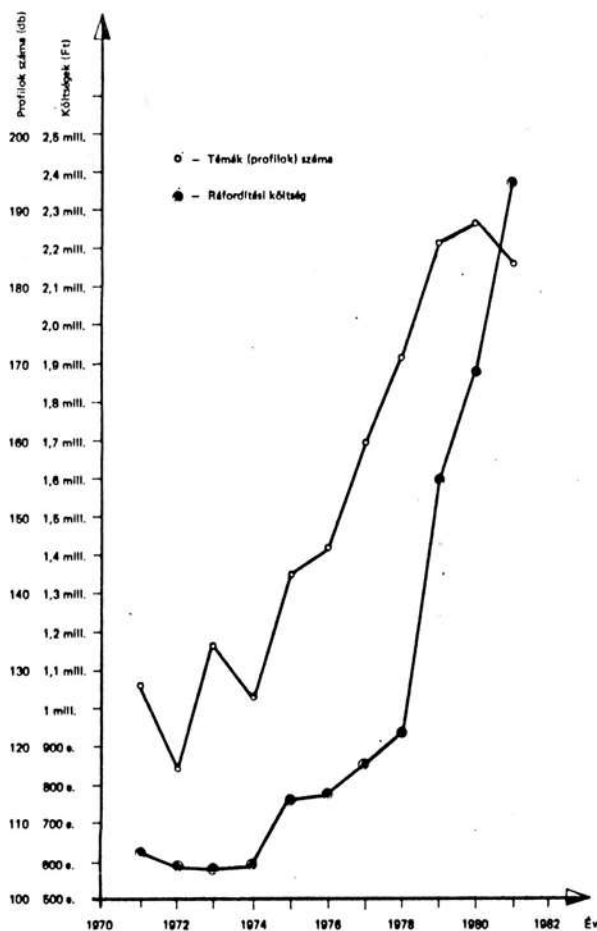
A Veszprémi Vegyipari Egyetem Központi Könyvtára a külső felhasználók számára nyújtott szolgáltatáshoz állami támogatást nem kap, a szolgáltatás díját pedig a jövőben sem kívánja emelni. A könyvtár célja a *költségek csökkentése*, úgy, hogy a szolgáltatás változatlan árszinten maradjon, ugyanakkor a feldolgozás színvonala, minősége ne változzék.

Költségcsökkentő tényezőként szóba jöhet:

a keresőprofilok felülvizsgálata, strukturális változtatása;

a visszakereső programrendszer módosítása.

A keresőprofilok felülvizsgálatával kapcsolatban azt a gyakorlatot alkalmazzuk, hogy az egy-egy témára írt profilokat legalább 4–5 alkalommal javítjuk, mielőtt rendszeres keresést végeznénk vele.

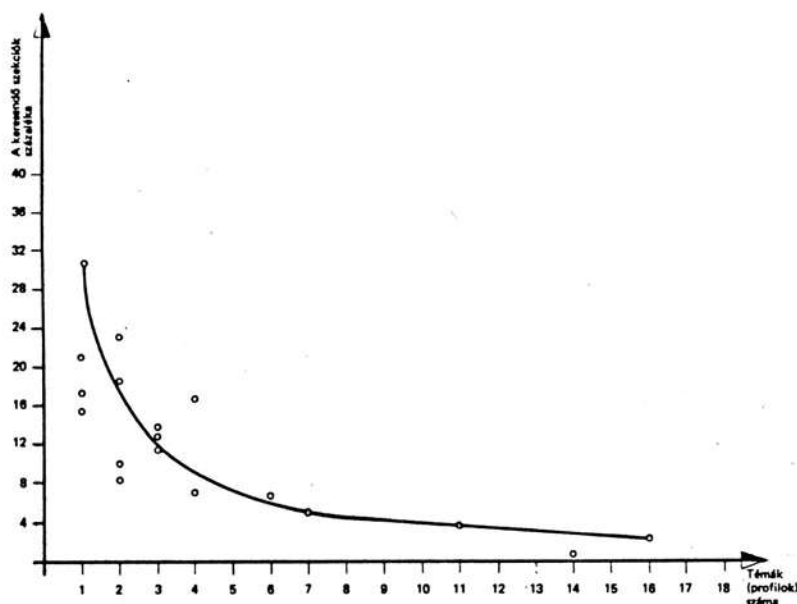


1. ábra A ráfordítási költségek és a témák (profilok) számának alakulása

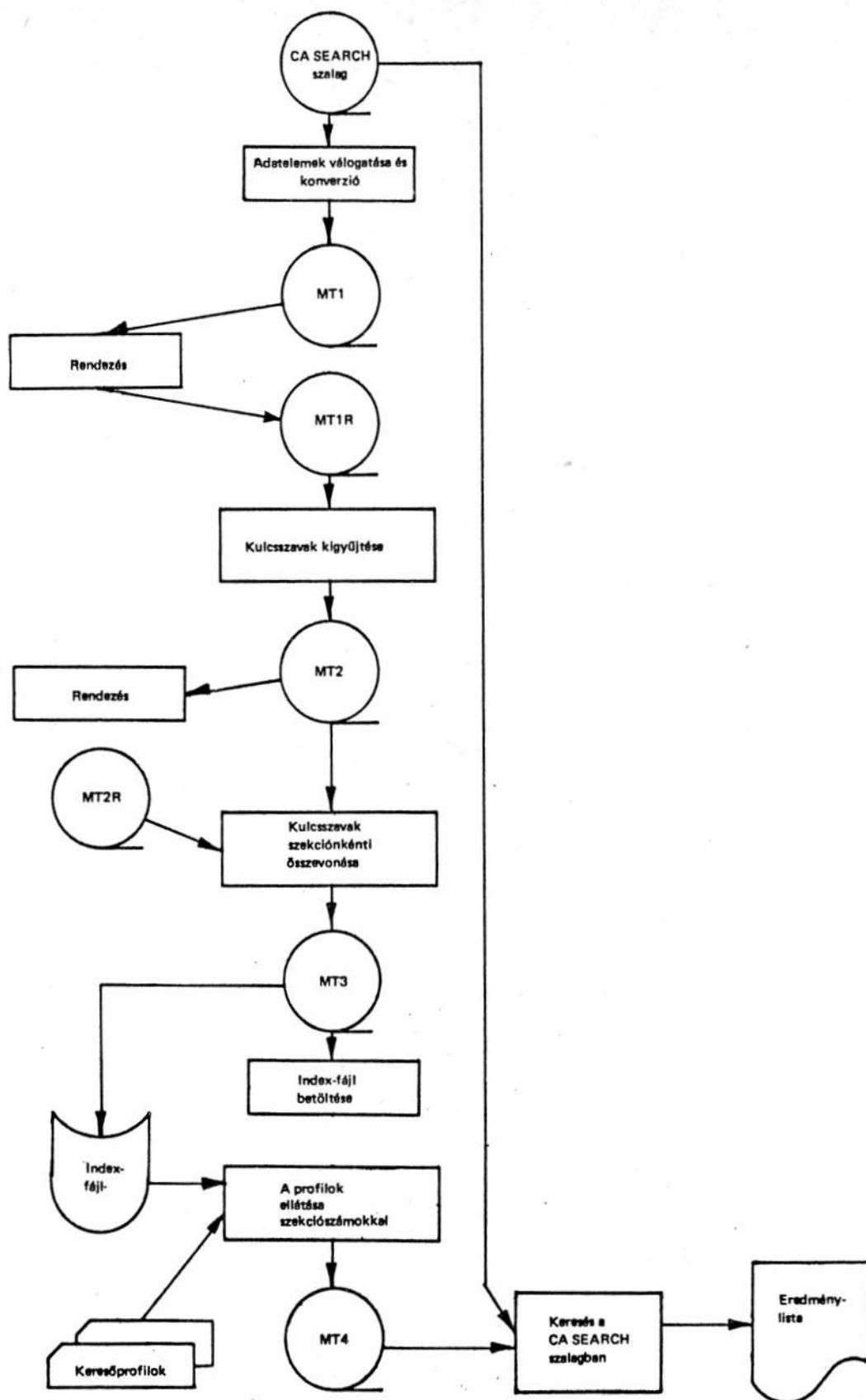
A keresőprofilban biztosított keresési lehetőségek korlátozásával, a keresőprofil strukturális változtatásával a keresés minőségét és megbízhatóságát rontanánk. Így csak a második lehetőség marad, a *visszakereső program-rendszer módosítása*.

A *Chemical Abstracts* és ennek mágnesszalagos formája, a CA SEARCH a logikai rekorodokat (dokumentumleírásokat) tartalmilag 80 szekcióba csoportosítja. Egy-egy témakörrel foglalkozó rekorodok csak bizonyos szekciókban fordulhatnak elő. Mivel a CAS azt az elvet követi, hogy egy dokumentumleírást csak egyszer közöl, és csak egy szekcióban helyezi el, gyakran nem lehet egyértelműen eldönteni, hogy egy bizonyos dokumentumleírás (főleg, ha határterülettel foglalkozik) melyik szekcióba kerülhet. Éppen ezért a gépi keresésnek nagy előnye az, hogy az összes rekodot átnézi. Ez a keresési módszer azonban a feldolgozást lassúvá teszi. Ezt az ellentmondást úgy tudjuk áthidalni, hogy *csak azokat a rekorodokat nézzük át, amelyek feltétlenül kielégítik a keresőprofilokat*.

A heti eredménylisták (a témakeresés során kapott output-jegyzékek alapján egy évig figyeltük, hogy az egyes keresőprofilokra a 80 szekcióból hány szekcióban találtunk válaszokat. Vizsgálatunk eredményét a 2. ábra mutatja. Az ábrából az tűnik ki, hogy a vizsgált 84 témából 48 téma esetén a 80 szekció alig 5%-át kell átnézni, azaz csak 1, 2, 3, ill. 4 szekciót. A fennmaradó 36 téma esetén pedig 5–25 szekciót kell átnézni. Ha előre meg tudnánk határozni a feldolgozáskor átnézendő szekciók számát, akkor sokkal gyorsabban történne a keresés.



2. ábra A keresendő szekciók százalékos alakulása a vizsgált témák (profilok) függvényében



3. ábra Előttét-programmal kiegészített visszakereső programrendszeren alapuló feldolgozás folyamatábrája

Ennek megoldására szolgál egy *előtét-program* beiktatása (3. ábra). Az előtét-program az első lépésben az input-szalagot rekordonként beolvassa, és azokat az adatelemeket válogatja ki, amelyekről egy ún. szótárt készít. Ebben a fázisban az ASCII → EBCDIC kód-konverziót is elvégzi. A következő fázisban az egyedi kulcsszavakat (egy rekordra nézve egyedi, mivel egy rekordon belül ugyanaz a kulcsszó többször is előfordulhat) rekordonként kiválogatja. Következő lépésben, a kulcsszavak szekciónkénti összevonásával először az MT2 fájl szekciónként rendezi, majd a rendezett fájlból (MT2R) az azonos szekciókban előforduló azonos kulcsszavakat kiszűri. Az így kapott index-fájlt (szótárt) mágneslemezre viszi, amelynek segítségével a keresőprofilokkal összevetve megállapítja, hogy a valamely keresőprofil kielégítő rekordokat melyik szekciókban kell keresni. Ezt a keresőprofilok számára feljegyzi, és ennek megfelelően korrigálja azokat.

A szekciós számokkal ellátott keresőprofilokkal azután az eredeti feldolgozó programrendszer segítségével a keresést végrehajtjuk. Az így kiegészített programrendszer üzemszerű használatával az előtét-programmal minden mágnesszalag esetén elkészítjük a szótárt, és ennek segítségével meghatározzuk az átnézésre kerülő szekciók számát.

A cikkben közölt kísérletet egy heti mágnesszalaggal (CA SEARCH 97/03) és a Veszprémi Vegyipari Egyetem Központi Könyvtaránál feldolgozásra kerülő keresőprofilok egyharmadával végeztük el. A kísérlet során az előtét-programmal kapott futási idők a következőképpen alakultak:

A szótár készítése egy mágnesszalagról:	371 s
A keresőprofilok ellátása szekciós számokkal:	267 s
A teljes előtét-program futási ideje:	638 s
Az eredeti keresőprogram futási ideje:	13 926 s
Az előtét-programmal korrigált (szekciókkal ellátott) keresőprofilok futási ideje:	12 360 s
A kétféle futtatásból származó időnyereség:	928 s

A kísérlet eredményeként a SYSTEM Szervezési Vállalat az új programrendszerrel 1983-ra a feldolgozást 200 ezer Ft-tal csökkentett összegért vállalta el. A Veszprémi Vegyipari Egyetem Központi Könyvtára a feldolgozásból származó megtakarítást a Chemical Abstracts Service-nek fizetendő profil-jogdíj kiegyenlítésére fordítja. Ily módon tudja a könyvtár a szolgáltatását a továbbiakban is változatlan árszinten biztosítani.

EGYHÁZY Tiborné—KERTÉSZ Miklós:
A CA SEARC mágnesszalagos feldolgozás
gyorsítása
Számítógépes Kémiai Szakirodalomfigyelő
Szolgálat a Veszprémi Vegyipari
Egyetemen

A cikk bemutatja a Veszprémi Vegyipari Egyetem Központi Könyvtárának a Chemical Abstracts Service CA SEARCH adatbázisára épülő szelektív információterjesztési szolgáltatását. A gépi keresés gyorsítására nemrégiben ún. előtét-program készült, amelynek beiktatása a keresési idő és így a feldolgozási költségek jelentős csökkenését eredményezte. E megtakarításnak köszönhető, hogy a Könyvtár a szolgáltatást továbbra is változatlan áron tudja nyújtani.

* * *

Mrs. EGYHÁZI, M.—KERTÉSZ, M.: Increasing
the processing speed of the CA SEARCH
database
The Computer-Assisted Chemical SDI Service
at the Veszprém Chemical University

The computer-assisted SDI service of the Central Library of the Veszprém Chemical University based on the CA SEARCH database of the Chemical Abstracts Service is described. In order to increase the computerized search a pre-processor program has been developed which has considerably reduced processing time and costs. As a result of this saving the Library could keep its SDI prices unchanged.

* * *

ЭДЬХАЗИ, Т. — КЕРТЕС, М.: Ускорение машинного поиска в CA SEARCH.

Служба информационного поиска на базе ЭВМ в веспремском Химическом университете

В работе представлена служба машинного ИРИ, осуществляемого на базе Chemical Abstracts Service CA SEARCH в центральной библиотеке веспремского Химического университета. С целью уменьшения времени машинного поиска недавно была создана так называемая подготовительная программа, которая обеспечила значительное уменьшение времени поиска и, следовательно, затрат по использованию машинного времени. Благодаря этой экономии библиотека и в дальнейшем может предоставлять эти услуги по неизменной цене.

Frau EGYHÁZY, M.—KERTÉSZ, M.:
Beschleunigung der Magnetbanddaten-
verarbeitung CA SEARCH
Rechnergestützter Chemischer
Fachliteraturbeobachtungsdienst an der
Universität für Chemieindustrie,
Veszprém

Im Artikel wird der selektive Dokumentationsdienst der Zentralbibliothek der Universität für Chemieindustrie, Veszprém dargestellt, das auf die Datenbasis CA SEARCH des Chemical Abstracts Service aufgebaut wurde. Zwecks Beschleunigung der maschinellen Recherche wurde vor kurzem ein sog. Vorsatzprogramm verfertigt, dessen Einfügung eine beachtliche Reduktion der Recherchezeit und demzufolge der Verarbeitungskosten bewirkte. Dank dieser Einsparungen ist die Bibliothek in der Lage, ihre Dienstleistungen zu unverändertem Preis anzubieten.

* * *

* * *

Az OKT alelnökének leveléből

1983. V. 19-én keltezett levelében Rózsa György a TMT 1983. 3. számában megjelent két közleményhez fűzött megjegyzést. Idézünk szerkesztőségünknek küldött leveléből:

„Wix Györgyné cikkében a 97. oldalon, valamint Sárdy Péter hozzászólásában a 100. oldalon a következő egybehangzó téves állítások szerepelnek: 'Azt is tudom, hogy az Országos Könyvtárügyi Tanács által létrehozott bizottság...' (Wixné); 'Maradt tehát a módszer, amelyet az OKT egy bizottsága...' (Sárdy).

Az OKT alelnökeként közlöm, hogy sem a Tanács, sem pedig a Tanács Elnöksége semmiféle bizottságot nem hozott létre, az akció lebonyolítása az OKT megkerülésével történt."

A szerkesztőség