

TÁJÉKOZTATÓ AZ ERICTAPES ADATTÁRRA ALAPOZOTT GÉPI INFORMÁCIÓKERESŐ RENDSZERRŐL

EGYHÁZY Tiborné – KOVÁTS Zoltán

Az információáradat minden szakterületen megnehezíti a szakemberek, kutatók, tudósok munkáját. Pedig a szakmai világirodalom módszeres áttekintése és tudatos válogatása nélkül kb. 50 %-os valószínűsége annak, hogy a kutatómunka csak fölösleges ismétlése egy régebben eredményesen befejezett kutatásnak, vagyis népgazdasági szinten a kutatásokra fordított több milliárdos nagyságrendű összeget a jelenleginél kétszerte jobb hatásokkal lehet felhasználni – megfelelő szakirodalmi előkészítés esetén.

A kutató „testére szabott” információk hézagtalan és gyors kiválogatását a minél teljesebb világirodalomból az elektronikus számítógép alkalmazása tette lehetővé. Magyarországon először a Veszprémi Vegyipari Egyetem Központi Könyvtára (VVEKK) szervezett számítógépes szakirodalomfigyelő szolgálatot a kémia tudományterületén. A szolgáltatás üzemeltetésében az 1971 óta szerzett tapasztalatokat felhasználva készült el az Országos Pedagógiai Könyvtár és Múzeum (OPKM) részére a pedagógiai információs rendszer korszerű, gépi információkereső szolgálatának terve.

A gépesített információs rendszer feltételei:

adattár

program

számítógép

információs szakember

a felhasználók szervezett együttese.

Adattár tekintetében az Egyesült Államok Pedagógiai Kutató Intézetének égisze alatt készülő ERICTAPES-re esett a választás.

BEVEZETÉS.

A pedagógia, oktatás, oktatás-technológia terén megjelent információk, kutatási eredmények közlésére és a felhasználók gyors tájékoztatására vállalkozott az Educational Resources Information Center (ERIC) of the National Institute of Education (NIE), amikor megjelentette az ERICTAPES mágnesszalagos adattárait. Az Intézet három fajta adattárat hoz forgalomba, a RESUME LINEAR FILE-t, a POSTINGS FILE-t, és a THESAURUS FILE-t. Ezen kívül nyomtatott kiadványai is ismereteseek: a Research in Education (RIE), a Current Index to Journals in Education (CIJE) stb.

Az ismertetett SDI (Selective Dissemination of Information) rendszer az első típusú file, a RESUME LINEAR FILE alapján működik. A RESUME LINEAR FILE kivonatokat tartalmaz növekvő számsorrendben a referáló laphoz hasonlóan. A RESUME LINEAR FILE-on belül a kivonatok típusának (folyóirat-cikk, riport kivonat), illetve a kétféle nyomtatott kiadványoknak megfelelően (RIE és CIJE) kétféle típusa van az adattáraknak is:

- REPORT RESUMES FILE
- JOURNAL RESUMES FILE.

AZ ADATTÁRAK FŐBB JELLEMZŐI

Az ERIC file-k 9 csatornás, páratlan paritású, nem címkézett mágnesszalagok. A file rekordjai az IBM S/360 Operating System (OS)-nek megfelelően változó blokkokban (VB) helyezkednek el. A rekordok sűrűsége a felhasználó igénye szerint 800 BPI vagy 1600 BPI. Visszakereső rendszerük az utóbbit, az 1600 BPI sűrűségű mágnesszalagot használja.

REPORT RESUMES FILE

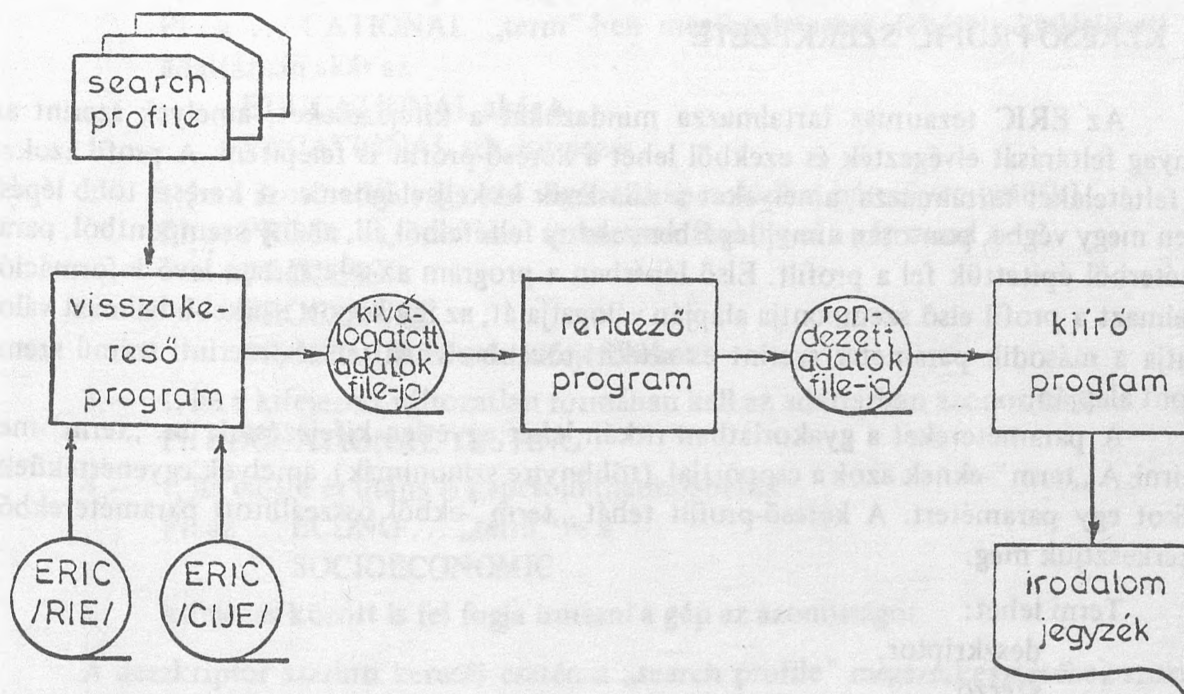
A Resources in Education (RIE)-ben közölt kutatási jelentések kivonatait tartalmazza. A RIE az oktatás területén megjelenő különféle kiadványokat ismerteti, rendszerezi és közli. A RIE-kivonatok értéke elsősorban abban rejlik, hogy a nem publikált, az úgynevezett „illó” irodalomról: technikai és kutatási jelentésekről, beszédekről, tanítási útmutatókról, statisztikai összeállításokról, tantervi anyagokról stb. ad tájékoztatást. A riport-kivonatok átlagos havi gyarapodása 1200 kivonat/hónap. A mágnesszalagok havi, negyedévi vagy éves egységekben rendelhetők meg. A mágnesszalagon levő riport-kivonat output formája az 1. számú mellékleten látható.

JOURNAL RESUMES FILE

A Current Index to Journals in Education (CIJE)-ban megjelenő cikk-kivonatokat tartalmazza. A CIJE közel 700 oktatási és oktatással kapcsolatos folyóiratból válogatja a referálandó közleményeket. A folyóiratcikk-kivonatok átlagos havi gyarapodása 1600 kivonat/hónap. A mágnesszalagok szintén havi, negyedévenkénti, vagy éves egységekben rendelhetők meg. A mágnesszalagon levő folyóiratcikk-kivonatok output formája a 2. számú mellékleten látható.

AZ ERIC TAPES-EN ALAPULÓ INFORMÁCIÓ-VISSZAKERESŐ RENDSZER ELVI VÁZLATA

Az információvisszakereső rendszer (1. ábra) szekvenciálisan (több lépésben, sorban haladva) működik. A mágnesszalagos adattárakból a program utasítására a számítógép a bennük sorban, egymás után elhelyezkedő kivonatokat egyenként behívja és összehasonlítja valamennyi „search profile”-al (kereső-profil). Ha a kivonat a „search profile”-ban megfogalmazott feltételeket kielégíti, akkor a gép egy külön mágnesszalagra felírja a kivonatot és mellé rendeli a neki megfelelő profilok azonosítóit. Ezután a rendező program a „kiválogatott” kivonatok file-jét sorba rendezi a „search profile”-ok azonosítója és a RIE, CIJE szerint. Az utolsó lépés a meghatározott formátumban, 2 példányban történő „listázás”.



1. ábra

A VISSZAKERESÉSHEZ SZÜKSÉGES, „SEARCH PROFILE” MEGSZERKESZTÉSE

A visszakeresés hatékonyságát elsődlegesen a „search profile” szerkezete befolyásolja. A „search profile” összeállításában aktívan részt kell vennie a felhasználónak, hiszen csak az adott területen dolgozó szakember ismerheti olyan mértékben a keresendő tárgykört, hogy az irodalomkereséshez szükséges kifejezések helyes összeválogatásával a legjobb eredményt adó „search profile”-t tudja előállítani a profile-készítő.

A visszakeresés az adattárban rögzített adattípusok szerint történhet, amelyeket a profile-ba beépítendő kifejezések előtt jelölni kell (betűkód):

RIE futtatásakor

- S szerzők szerinti keresés
- D deskriptorok szerinti keresés

CIJE futtatásakor

- S szerzők szerinti keresés
- D deskriptorok szerinti keresés
- C folyóirat szerinti keresés
- A azonosítók szerinti keresés.

A KERESŐ-PROFIL SZERKEZETE

Az ERIC tezaurusz tartalmazza mindazokat a kifejezéseket, amelyek szerint az anyag feltárását elvégezték és ezekből lehet a kereső-profil is felépíteni. A profil azokat a feltételeket tartalmazza, amelyeket a válasznak ki kell elégítenie. A keresés több lépésben megy végbe, pontosan annyi lépésben, ahány feltételből áll, ahány szempontból, paraméterből építettük fel a profilt. Első lépésben a program az adattárban levő információhalmazt a profil első szempontja alapján válogatja át, az így kapott szűkebb halmazt válogatja a második paraméter szerint és szűkíti tovább elvben tetszés szerinti számú szempont alapján.

A paramétereket a gyakorlatban ritkán lehet egyetlen kifejezéssel, ún. „term”-mel leírni. A „term”-eknek azok a csoportjai, (többnyire szinonimák), amelyek egyenértékűek, alkot egy paramétert. A kereső-profil tehát „term”-ekből összeállított paraméterekből szerkesztjük meg.

Term lehet:

- deskriptor,
- szerző,
- folyóirat címe,
- azonosító,

attól függően, hogy az adott témakörben a felhasználó melyik adatelem szerint kerestet. A term-ek között:

OR = VAGY

NOT = NEM

logikai kapcsolat állhat.

Paraméter: a hasonló, szinonim, logikailag egybetartozó term-ek egy csoportja. A paraméterek között:

AND = ÉS

logikai elem áll. A tiltó logikai elemmel kapcsolt term-et csak az első paraméterbe lehet tenni.

A „term”-ek megadása négyféle „mód”-on történhet, ami azt a célt szolgálja, hogy a számítógép a profile-ban megadott szótöredéket is felismerje az adattárban rögzített adatelemek között.

A problémát az jelenti, hogy a dokumentum címében, a teauruszban előforduló kifejezések is ragozott, toldalékolt formában fordulnak elő, a gép ilyen esetekben csak akkor ismerné fel a „term”-eket, ha vagy minden toldalékolt alakját is beépítenénk a paraméterbe, vagy ha a toldalékok bevágásával a főalakokat állítjuk elő. Más módozatokhoz hasonlóan az ERIC keresésére írt program is úgy oldja meg ezt a problémát, hogy lehetővé teszi csonkolt formák azonosítását. Az alábbi példák azt mutatják be, hogy ebben az esetben milyen utasításokat lehet eszközölni.

1 – jelet kell alkalmaznunk a keresőprofilba beépített „term”-ek előtt, ha a szótöredék vagy szó elejéhez kapcsolódhat bármilyen betű(k).

Pl. a ... CATIONAL „term”-ben megfogalmazott feltételt kielégítheti az adattárban akár az

EDUCATIONAL akár a
VOCATIONAL stb. kifejezés.

2 – t, ha a szótöredék vagy szó végéhez kapcsolódhat bármilyen betű(k)

Pl. a PROC... „term”-mel azonosnak fogadja el a gép mind a

PROCESS
PROCEDURE
PROCEDURES stb. kifejezéseket.

3 – t, ha a kifejezést változatlan formában kell az adattárban azonosítani.

Pl. EDUCATIONAL TESTING

4 – t, ha előtte és utána is kapcsolódhatnak betűk.

Pl. az ... ECONO ... „term” és a
SOCIOECONOMIC

kifejezés között is fel fogja ismerni a gép az azonosságot.

A deskriptor szerinti keresés esetén a „search profile” megszerkesztéséhez szükséges deskriptorokat az aktuális (legfrissebb) THESAURUS-ból, szabványosított tárgyszógyűjteményből kell kiválasztani és a logikai elemek segítségével a „search profile”-lá összeépíteni. (Egy profilon belül többféle adatelem is szerepelhet!)

„Természettudományos nevelés” című témakörben a „term”-ek és paraméterek kapcsolata a következő lehet:

1. Paraméter

EDUCATIO (term)
VAGY
INSTRUCTIO (term)

ÉS

2. Paraméter

SCIENC (term)
VAGY
BIOLOG (term)
VAGY
CHEMIST (term)
VAGY
PHYSIC (term)

A fenti kereső-profil alapján mindazokat a kivonatokat találatnak minősíti a gép, amelyeknek a deskriptorai között a következő kettő együttesen fordul elő.

Így:

EDUCATIO	+	SCIENC
EDUCATIO	+	BIOLOG
EDUCATIO	+	CHEMIST
EDUCATIO	+	PHYSIC
INSTRUCTIO	+	SCIENC
INSTRUCTIO	+	BIOLOG
INSTRUCTIO	+	CHEMIST
INSTRUCTIO	+	PHYSIC

„Term”-ek súlyozása.

A kereső-profil szelektivitását a term-ek súlyozásával lehet fokozni. Az egyes term-ekhez rendelt súlyértékek (–999-től +999-ig) a term-ekhez AND = ÉS logikai elemmel kapcsolódnak.

A keresőprogram biztosítja a kereső-profile-ban meghatározott és az adattárban felismert valamennyi term súlyértékének összegezését egy adott kivonaton belül. Ha az összeg egy előzetesen meghatározott küszöbérték alatt marad, a gép nem írja ki az egyébként találatnak tekinthető adatot.

Ezt képlettel a következőképpen lehet kifejezni:

$$\sum \text{Azonosított term-ek súlya (egy kivonaton belül)} \geq \text{küszöbérték}$$

Az előző példa „search profile”-ja, a súlyozás következtében az alábbiakban módosulhat:

Küszöbérték: 10

1. Paraméter

termek súlyértéke

2. Paraméter

termek súlyértéke

4	ÉS	EDUCATIO	}	ÉS	}	3	ÉS	SCIENC
		VAGY				3	ÉS	BIOLOG
4	ÉS	INSTRUCTIO				3	ÉS	CHEMIST
								VAGY
						3	ÉS	PHYSIC

Tehát a gép az előbbiek szerint csak azokat a kivonatokat fogja találatnak minősíteni és kiírni, amelyben a következő deskriptorok együttesen fordulnak elő:

4	EDUCATIO	+	3	SCIENCE	+	3	BIOLOG	=	10
4	EDUCATIO	+	3	SCIENCE	+	3	CHEMIST	=	10
4	INSTRUCTIO	+	3	SCIENCE	+	3	BIOLOG	=	10
4	INSTRUCTIO	+	3	INSTRUCTIO	+	3	SCIENCE	>	10

Tájékoztatónkban elsősorban az adattár szerkezetére, a program működésének főbb vonásaira, főleg pedig a keresőprofilok szerkezetének technikájára vonatkozó ismereteket kívántuk röviden összefoglalni.

A program készítésének pénzügyi fedezetét a Kulturális Minisztérium Könyvtárügyi Osztálya biztosította, megszerkesztésének feladatát a Számítástechnikai és Ügyvitelszervezési Vállalat (SZÜV) zalaegerszegi szakemberei vállalták. A számítógépet ugyancsak a SZÜV zalaegerszegi részlege biztosítja.

A felhasználók felkutatása és megszervezése, valamint a szolgáltatás eredményes üzemeltetése már az OPKM feladata.

1. melléklet

RIE KIVONAT OUTPUT FORMÁJA

- 1 – a futás ideje
- 2 – lapszámozás
- 3 – profil-azonosító
- 4 – a nyomtatott forrás megjelölés
- 5 – a profilon belüli lapszámozás
- 6 – a kivonat sorszáma
- 7 – a nyilvántartásba-vétel és egyéb technikai jellegű műveletek dátuma
- 8 – témacsoport
- 9 – a lehetséges témacsoportok
- 10 – a publikáció típusa
- 11 – programterület
- 12 – megjelenés éve
- 13 – cím
- 14 – szerző(k)
- 15 – deskriptorok
- 16 – azonosítók
- 17 – EDRS – Eric Documentation Reproduction Service (az intézmény reprográfiai szolgálatától megrendelhető xerox másolat vagy microfishe ára)
Jelen esetben EDRS-től a dokumentum nem szerezhető be.
- 18 – Terjedelem
- 19 – RIE megjelenési ideje
- 20 – kivonat
- 21 – a témát kidolgozó intézmény neve és címe
- 22 – Felügyeleti Szerv
- 23 – a beszerzéshez szükséges adatokat közli

CIJE KIVONAT OUTPUT FORMÁJA

- 1 – a futás ideje
- 2 – profilazonosító
- 3 – a nyomtatott forrás megjelölése
- 4 – a profilon belüli lapszámzás
- 5 – a kivonat sorszáma
- 6 – a nyilvántartásba vétel és egyéb technikai műveletek időpontja
- 7 – témacsoport
- 8 – a folyóirat megjelenési éve
- 9 – a cikk címe
- 10 – szerző(k)
- 11 – deszkriptorok
- 12 – azonosítók
- 13 – CIJE megjelenési időpontja
- 14 – kivonat
- 15 – a folyóirat címe, bibliográfiai adatai

RANGANATHAN DÍJ elnyerésére hirdetett pályázatot, immár második alkalommal, a FID Osztályozáskutatási Bizottsága (Comitee on Classification Research). A pályamunkák elbírálására különbizottságot hozott létre, amely megfelelő pályamű hiánya esetén fenntartja a jogot nemleges döntésre is. A díjat 1978 szeptemberében a FID 39. konferenciáján nyújtják át Edinburghban (UK). Az első Ranganathan díjat Mr. Derek AUSTIN kapta (British Library).

(A FID/CR körlevélből)