

Befejezésül, alapul véve a hazai mikrofotográfia jelenlegi helyzetét, felszereltségét és gazdasági alapjait, dönteni kell arról, hogy az egyfokozatú eljárásokon túlmenően milyen mértékben fordíthatunk figyelmet a

kétfokozatú mikrofilmtechnika bevezetésének előkészítésére.

Rajnák Antal

REFERÁTUMOK

Programcsomag kiválasztása számítógépes információszolgáltató rendszerhez

A feladat kitűzése

Számítógépes információszolgáltató rendszer szervezésekor nagyon munkaigényes dolog saját programrendszer kifejlesztését. Ezért terjed az a gyakorlat, hogy ilyen célra az elérhető kész programcsomagok közül választanak. Ezek hatékonysága nagymértékben függ attól, hogy mennyire illeszkednek az adott konkrét feladathoz, de kiválasztásukhoz nincsenek megbízható módszerek. Mivel a kész programrendszerek leírása általában nem a felhasználóknak szól, hanem a programozóknak, rendszerint ők döntenek a kiválasztásról, számukra viszont fontosabb szempont a könnyű üzemeltethetőség, mint például a keresés hatékonysága.

Az elmondottak alapján a VINITI *Információs keresőrendszerek* osztálya a következő általános feladatot tűzte maga elé:

általánosítani kell a tapasztalatokat, amelyek a kész programcsomagokkal működő információs keresőrendszerek tervezése során felgyűltek;

ki kell dolgozni a programcsomagok egységes leírásának rendszerét, megkönnyítve összehasonlításukat és megismerésüket, a rájuk alapozott rendszer tervezését; el kell készíteni a számbajövő programcsomagok szerkezetének és képességeinek leírását;

ki kell dolgozni a programcsomagok konkrét esetre vonatkozó értékelésének és kiválasztásának módszerét.

A fenti feladat megoldását *konkrét példán mutatjuk be*, hangsúlyozva, hogy a leírtakat nem tekinthetjük általános receptnek. A programcsomagok értékelése csak az adott esetre alkalmazható. A kiválasztott eset feltételrendszere a következő:

nagyméretű információs központ, széles tematikájú információs szolgáltatással;

az igényelt szolgáltatások: szelektív információterjesztés, retrospektív irodalomkutatás, párbeszédes információkeresés;

a feldolgozandó adatbázisok: a VINITI-ben készített valamennyi (bibliográfiai és faktográfiai) adatbázis;

a számbajövő programcsomagok: USS [1], ASOD [2], Poisk-1 [3].

A vizsgált programcsomagok

A programcsomagok vizsgálatát a következő szempontokra alapoztuk: az adatok bevitele, módosítása, tárolási módja, a keresőkérdés bevitele, keresési mód, keresési eredmény, diagnosztikai és statisztikai információk kiírása, a párbeszédes üzemmód során a felhasználóknak nyújtott segédinformációk.

Az egyes programcsomagoknak a számítógép-konfiguráció iránti igényével részleteiben nem foglalkoztunk. A konfiguráció-igény csak akkor befolyásolja a döntést, ha valamelyik programcsomag futtatásához nem elegendő a rendelkezésünkre álló lehetőség (pl. a *Poisk-1* nem futtat, ha a központi tár 256 kB alatt van). Esetünkben mindhárom vizsgált programcsomag igényeit kielégítik a rendelkezésre álló gépi konfigurációk.

USS (Unificated Software System – Egységesített szoftver-rendszer)

Szekvenciális keresésre, kötegelt feldolgozásra orientált rendszer. Jellegzetessége a formátum-egységesítés. Az adatbázis és a keresőprofilok formátuma a makroszintaxis szintjén azonos, a mikroszintaxis szintjén azonban már nem.

A keresőprofilok felépítése bonyolult, hosszas előkészítést igényel. Sokféle lexikai egységet használ (szó, szótöredék, szócsoport, szám stb.); ezeknek a szövegheli

helyzetét azonosítók írják le. Jobb- és baloldali csonkolás, maszkolás lehetséges. Az elemi keresési műveletek közé tartozik az egybeesés (egyenlőségen) kívül a „kisebb” és „nagyobb” reláció.* A végső kiválasztási kritériumnak az elemi keresési műveletekből történő kialakításában nemcsak a Boole-műveletek alkalmazhatók, hanem az aritmetikai összeadás és szorzás is. A keresőprofilra vonatkozó korlátozások csekélyek. Az output részletességére vonatkozó szabályozási lehetőség kevés: a dokumentum két részre osztható, ezek egyikét vagy mindkettőt lehet az outputra kérni.

A kereső operátorok gazdag választéka ellenére hiányoznak a szabadszöveges keresésben (különösen az orosz nyelvű keresésben) nélkülözhetetlen pozíciós operátorok, így a szavak egymáshoz viszonyított elhelyezkedése nem vehető figyelembe.

A keresés kis adatmennyiségben (szelektív információterjesztés) elég gyors, nagy adatmennyiség esetén (retrospektív keresés) azonban a szekvenciális keresési mód miatt annyira időigényes, hogy ilyen célra az USS nem ajánlható.**

*АСОД (Автоматизирующая Система
Обработки Документов – Автоматизированный
документумфедолгозый rendszer)*

Fő jellegzetessége a szövegkezelési lehetőségek bősége. Az információt adatbázisokba csoportosítva tárolja. Az adatbázisok soros állományban őrzött formatizált mezőkből és invertált állományban őrzött szövegrészből állnak. Mind az adatbevitelnél, mind a keresésnél nagyon sokoldalú szövegkezelési lehetőségei vannak: keresésből kizárt szavak kijelölése, szinonimakizárás, szavak helyettesítése stb. Kétféle összefüggést tartalmazó teauruszt tud kezelni.

A mezőkre, mondatokra, szavakra tagolódozó szöveges részben csonkolás nélkül, jobboldali csonkolással, teauruszhasználattal lehet keresni. Háromféle pozíciós operátor is használható: egy mezőbe, egy mondatba, meghatározott szónyi távolságon belülré esés. A formatizált részben egybeesésre, különbözőségekre, töredék előfordulására, maszkolt egybeesésre és aritmetikai relációkra lehet keresni. Az elemi keresési műveletek eredménye között a szöveges részben VAGY, ÉS és NEM, a formatizált részben VAGY és ÉS kapcsolat írható elő, a két rész keresése egymással ÉS kapcsolatban van. A keresésben súlyozás is előírható. Szabadon írható

* A karakteres összehasonlítás értelmében is. – *A ref. megj.*

** Magában az USS-t kidolgozó intézetben, a prágai UVTEI-ben sem ezt használják retrospektív keresésre, hanem a GOLEM-et. – *A ref. megj.*

elő, hogy az outputban a dokumentum mely mezői jelenjenek meg. A keresésről bőszeges statisztikai információt ad.

A programcsomag párbeszédés üzemmódra is rendelkezésre áll.

A keresőprofil méretére és a formatizált mezők számára vonatkozó korlátozások faktografikus keresésnél zavaróak lehetnek.

Poisk-1

Ennek a programcsomagnak különösen fejlettek a párbeszédés lehetőségei. Az információ itt is adatbázisokra, az adatbázisok szorosán tárolt formatizált mezőkre és invertált állományban tárolt szöveges részre tagozódnak. A formatizált mezők itt meghatározott hosszúságúak, azonosítóval rendelkeznek. A szöveges rész mezőkre, mondatokra, szavakra tagolódik, a mezőknek azonosítójuk és hívójelük van. A szövegkezelés szűk: bizonyos szavak illetve mezők kizárására korlátozódik. Csak egyféle összefüggéssel rendelkező teauruszt kezel.

A szöveges részben történő keresésnél a tetszőleges jobboldali csonkoláson kívül megadott betűszámra vonatkozó csonkolás is lehetséges. Teaurusz a keresésben használható. Csak „durva” pozíciós operátorok vannak: egy mezőbe, egy mondatba, meghatározott nevű mezőbe esés. A „finom” pozíciós (kontextuális) operátorok hiányoznak, így szókapcsolatok keresése nehézkes.

A formatizált részben egybeesésre, különbözőségekre, maszkolt egybeesésre és aritmetikai relációkra lehet keresni, de töredékre nem. A szöveges részben VAGY, ÉS és NEM kapcsolatba, a formatizált részben VAGY és ÉS kapcsolatba hozhatók az elemi keresési eredmények.

Öt output-formátum között lehet választani, és az output a formatizált mezők valamelyike szerint rendezhető. Bőszeges statisztikai és kiegészítő információt nyújt a keresésről.

Különleges szolgáltatásai a programcsomagnak a több (legfeljebb 16) hasonló szerkezetű adatbázisban egyidejűleg végezhető keresés, a keresőprofilok tárolhatósága és visszahívhatósága egy újabb keresésben, a kötegelt feldolgozás lehetősége.

A programcsomag főbb hibái a „finom” pozíciós operátorok már említett hiányán kívül a szövegkezelés korlátozottsága, a formatizált mezők kötött hossza és legfőképpen a formatizált rész korlátozott terjedelme, amely lehetetlenné teszi a faktografikus keresést.

A három programcsomag főbb jellemzőit az 1. táblázat tartalmazza.

Az USS, ASOD és Poisk-1 programcsomagok főbb jellemzői

Program-csomag	Szöveg-szerkesztés az adatok bevitelekor	Szelektív információ-terjesztés lehetősége	Retrospektív keresés lehetősége	Párbeszédés üzemmód	Tezaurusz-használat a keresésben	Faktografikus keresés	Tetszőleges töredékre történő keresés	Pozíciós operátorok	Output-vezérlés
USS	nincs	van	nincs	nincs	nincs	jól alkalmazható	lehetséges	nem használható	korlátozott
ASOD	sokoldalú	van	van	van	három* kapcsolattípussal	jól alkalmazható	lehetséges	„finom” és „durva”	sokoldalú
Poisk-1	korlátozott	van	van	fejlett	egy kapcsolattípussal	korlátozott	nincs	csak „durva”	öt formátum

* A szövegben két kapcsolattípusról ír a szerző, itt tehát ellentmondás van. – *A ref. megj.*

A feladat megfogalmazása

Feladatunk elemei általánosságban a következők:

b_1, \dots, b_n a feldolgozandó adatbázisok;

$f_1^i, \dots, f_{H(i)}^i$, ahol $i = 1, \dots, n$, a megvalósítandó szolgáltatások, $H(i)$ az i -edik adatbázissal megvalósítandó szolgáltatások száma;

S_1, \dots, S_m a választható programcsomagok.

Esetünkben:

$b_1 = BK$ típusú adatbázisok (bibliografikus danye i klüचेve slova – bibliográfiai adatok és kulcsszavak);

$b_2 = SD$ típusú adatbázisok (strukturnye danye – szerkezeti adatok), amelyek a bibliográfiai adatokon és a kulcsszavakon kívül a kémiai vegyületek szerkezetére vonatkozó adatokat is tartalmaznak;

f_1^1 és $f_2^1 = a$ BK típusban történő bibliográfiai szelektív információterjesztés illetve retrospektív keresés;

f_1^2 és $f_2^2 = az$ SD típusban történő bibliográfiai szelektív információterjesztés illetve retrospektív keresés;

f_3^2 és $f_4^2 = az$ SD típusban történő faktografiai szelektív információterjesztés illetve retrospektív keresés;

$S_1 = USS$;

$S_2 = ASOD$;

$S_3 = Poisk-1$.

Az S_1, S_2, S_3 programcsomagok közül úgy kell egyet vagy többet kiválasztanunk, hogy az f_1^1, \dots, f_4^2 feladatok mindegyike megoldható legyen. A megoldásnál figyelembe kell venni, hogy:

a választás hosszú időre meghatározza a megkezdett munkát;

a megindítandó szolgáltatásnak a döntést meghatározó körülményei később megváltozhatnak;

a döntésnél nagyon sok szempontra kell egyidejűleg tekintettel lenni, ezek ellentétes követelményeket támaszthatnak;

a programcsomagokról nem áll rendelkezésünkre pontos és teljes információ – döntenünk kell, mielőtt igazán megismerhetnénk őket.

Az összehasonlító elemzés csak döntéselőkészítés. A tulajdonképpeni döntés egy felelős személy feladata. Ezért az elemzésben a döntésre jogosult személy szempontjait kell érvényesíteni, a kritériumokat az ő szempontjai szerint kell kiválasztani és súlyozni. A szempontokat értékelő skálák csak kvalitatívak lehetnek.

A döntés vázlata

1. szakasz. Olyan $\gamma^1, \gamma^2, \dots, \gamma^m$ döntési szempontokat és hozzájuk tartozó értékelési skálákat dolgozunk ki, amelyekkel a $\langle b_i, f_h^i, S_j \rangle$ hármassokat egyenként vizsgálhatjuk. Ha van olyan S_j , amely ideálisan teljesíti a feltételeket, azt választjuk. Ellenkező esetben kiküszöböljük azokat a hármassokat, amelyek a döntésre jogosult személy számára elfogadhatatlanok. Végül minden $\langle b_j, f_h^i \rangle$ párhoz csak egyetlen $\langle b_i, f_h^i, S_j \rangle$ hármast tartunk meg.

2. szakasz. A megmaradt hármassokat $\Gamma^1, \Gamma^2, \dots, \Gamma^m$ általános rendszerszempontok és a hozzájuk tartozó értékelő skálák szerint vizsgáljuk a döntésre jogosult személy által megadott döntési szabály szerint. Ha nem sikerül döntenie, akkor visszatérünk az 1. szakaszhoz, módosítjuk a $\langle b_i, f_h^i, S_j \rangle$ hármassok szelektálását.

Előzetes elemzés

Két alapgazdálás közül kell választanunk. Ha az egy adatbázison végzett többféle kereséshez megengedjük más-más programcsomag használatát, akkor maximális hatékonysággal oldhatunk meg minden feladatot. Ha viszont egy adatbázishoz csak egy programcsomagot engedünk meg, nem kell egy adatbázist többször konvertálnunk, és nem kell a személyzetet több programcsomag használatára betanítanunk, jobban üzemelhet a rendszer.

A második megoldást választottuk, így a retrospektív keresésre alkalmatlan USS kiesett a versenyből. Egyszerűsége mind a $\langle b_i, f_h^i, S_j \rangle$ hármasok vizsgálatát a $\langle b_i, S_j \rangle$ párok vizsgálatára egyszerűsítettük.

Értékelési szempontok

Az ASOD és a Poisk-1 összehasonlítása során elhagytuk az egyelőre nem értékelhető szempontokat. A két értékelési szakaszra a következő szempontokat állítottuk össze, minden szempontra két- vagy háromfokozatú értékelési skálával.

1. szakasz

- K1 – az összes keresési feladat megoldhatósága;
- K2 – szövegkezelési lehetőségek;
- K3 – a párbeszédés üzemmód lehetőségei;
- K4 – az üzemeltetés megindítására való felkészültség (a szoftver és a dokumentáció készlete, az üzemeltetési tapasztalatok szintje).

2. szakasz

- KI – az információs rendszer valamennyi feladatának megoldhatósága.
- KII – üzemeltetési szempontok. (Egy programcsomag üzemeltetése elegendő, vagy mindkettőre szükség van.);
- KIII – a párbeszédés üzemmód lehetősége valamennyi adatbázisra.

Az adatbázisok tulajdonságai és azok hatása a választásra

Az eddig BK típusúnak nevezett adatbázisok két csoportra oszthatók. Egyik csoportjukban (nevezük BK₁-nek) döntő jelentősége van a szótöredékre irányuló keresésnek. Ilyenek például a kémiai, matematikai, fizikai adatbázisok. Vegyük példának a kémiát. A poliamidok csak úgy kereshetők, ha a poli szótöredéket a szó kezdeteként írjuk elő, az amid szótöredék pedig a szó bármely részén előfordulhat. Ha a két töredéket külön-külön nem tudjuk keresni, a poliamidokat nem kereshetjük, hiszen valamennyit felsorolni lehetetlen.

A többi BK típusú adatbázisban (jelöljük őket BK₂-vel), például az informatikában, nincs jelentősége a szótöredékre keresésnek.

Az SD típusú adatbázisokban szintén döntő fontosságú a szótöredékre irányuló keresés. Ezenkívül a szótöredék elérheti azoknak a mezőknek a számát, amelyekben aritmetikai relációra kell keresni. A Poisk-1 csak 25 ilyen enged meg.

Értékelés és döntés

A döntés 1. szakasza során egyik programcsomag sem bizonyult ideálisnak, csak kompromisszumok árán lehet elfogadni bármelyiket is.

A BK₁ és SD típusú adatbázisok esetében döntő a K1 kritérium: a Poisk-1 nem alkalmas valamennyi feladat megoldására. A BK₂ típusú adatbázisok esetében a K2 kritérium az ASOD, a K3 kritérium a Poisk-1 mellett szól. A K4 kritérium minden információs központra külön értékelendő.

A 2. szakaszban így két változat közül kell választanunk:

a) az ASOD kizárólagos használata;

b) az ASOD használata a BK₁ és SD, a Poisk-1 használata a BK₂ típusú adatbázisokra. A Poisk-1 kizárólagos használatát a KI kritérium kizárja. A KII kritérium az a) változat, a KIII kritérium a b) változat mellett szól. A döntésre jogosult személynek aszerint kell döntenie, hogy ő a KII illetve a KIII szempontok közül melyiket tartja fontosabbnak.

Összefoglalás

Számítógépes információs szolgáltató rendszer szervezésekor általában kész programcsomagok közül kell választani. Ezek között a választás nehéz. Sok szempontot kell figyelembe venni, és az értékelés csak kvalitatív lehet. A feladat megoldására kétlépéses döntési séma javasolható: az első lépésben minden adatbázishoz külön-külön választjuk ki a legmegfelelőbb programcsomagot, a második lépésben rendszerszinten döntünk.

A javasolt séma jelentősen megkönnyíti a választást, és a döntésre jogosult személy számára lehetőséget nyújt a megalapozott döntésre. A javasolt módszer megfelelő módosításokkal alkalmassá tehető más és hasonló feladatok megoldására is.

Irodalom

1. Návod na jednotnou formulaci rešeršních dotazů v systému Ústřední technické základni. (Verse 1.) Prága, SÚTZ/FD – Dubeň, 1974.
2. Metodické materiály po primeneniu paketa prikladnyh program ASOD. Časti 1–4–M. Moszkva, NTMIK, 1978.

3. Paket prikladnyh program „Poisk-1”. Rukovodstvo polzovatelja. IP-011-5-16-78-1. Rukovodstvo programista. IP-011-3-6-78-1. Szófia, Interprogramma, 1978.

/LEONTEVA, T. M. – KRASIL'SIK, I. SZ. – ROZEMAN, M. I.: O zadačah vybora paketov prikladnyh programm dlja realizacii informacionno-poiskovyh sistem. = Naučno-Tehničeskâ Informaciâ, 2. sor. 4. sz. 1981. p. 10–15./

(Válasz György)

AZ INSPEC számítógépes dokumentum-beszerezési rendszere

A nagy és jól ismert brit referáló szolgáltatási és adatbázis-előállító intézmény, az INSPEC, új online dokumentumbeszerezési és nyilvántartási rendszert helyezett üzembe. Ezentúl ez szolgálja a fizikai, elektronikai, automatikai és számítástechnikai adatbázis nyersanyagául szolgáló primer dokumentumok beszerzésének jobb kézben tartását, ellenőrzését.

Az adatbázis mintegy 2600 folyóirat, valamint sokezer konferencia-anyag, kutatási jelentés, szakkönyv és disszertáció feldolgozására épül. 1981 januárja előtt ezek beszerzését, adminisztrációját és nyilvántartását hagyományos könyvtári módszerekkel végezték. E munka minden funkcióját egy PDP 11/70 típusú miniszámítógépre alapozott online rendszerbe szervezték át.

Az új rendszer fő funkciói a következők:

- az új anyagok bibliográfiai és egyéb adatainak felvétele,
- a már korábban bevitt dokumentumok adatainak kikeresése ill. módosítása,
- a hiányzó vagy késve érkező kiadványok jelzése,
- statisztikai táblázatok készítése.

A rendszerben feldolgozandó kiadványok (folyóiratok, könyvek, reportok stb.) adatairól egy-egy rekord készül, mégpedig külön-külön a megrendelésre javasolt és a megrendelt, de még be nem érkezett kiadványokról. A folyóiratok rekordjai további alrekordokat is tartalmaznak, a bennük publikált önálló cikkekről, közleményekről. Ezeknek a rekordoknak az összességéből áll a gépi rendszer fő fájlja.

Mielőtt egy-egy dokumentum adatait beviszik a fő fájlba, meg kell győződni arról, hogy nincs-e már rekordja. Ugyanis ez teszi lehetővé egyrészt a régi rekord aktualizálását (pl. egy folyóirat legfrissebb számának bekerülése, névváltozás stb.), másrészt ez nyújt védelmet a duplikálás ellen. Az összehasonlításra használt legegyszerűbb „kulcsok”: időszaki kiadványokra az ISSN, nem-időszaki kiadványokra az ISBN, s ha ezek

hiányoznak, akkor minden kiadványnak van egy meghatározott hatbetűs kódja.

Ha a rekord itt még nem szerepel, akkor az adatok bevitele után a fájl az új rekorddal gyarapodik. Ekkor a rekord automatikusan kap egy belső azonosítószámot; ezzel lehet a továbbiakban – feldolgozás vagy aktualizálás céljából – a legegyszerűbben hozzáférni a kívánt rekordhoz.

A fő fájl adatait képernyőn lehet lekérdezni, és billentyűvel bevinni. Az online rendszer a kezelőt egyszerű módon vezeti akár a kívánt dokumentum kikereséséhez, akár az új rekord helyes bevételéhez. Az adatrögzítési hibákat a rendszer automatikusan ellenőrzi, és csak hibátlan adatokat fogad el.

A forrásdokumentum rekordjának adatait egyesítik annak egyes részdokumentumaihoz rendelt alrekordok adataival. Tehát egy folyóiratcikk bevételéhez elegendő csak a cikk adatait beírni, ezeket a rendszer automatikusan kiegészíti a folyóiratnak mint kiadványnak az adataival. Visszakereséskor már egységes, teljes rekord áll a feldolgozók rendelkezésére.

A beszerzési rendszer igen jól használható a hiányzó vagy késésben lévő számok kijelzésére. A számítógép a korábbi érkeztetési adatokból kiszámítja, hogy mikorra várható a következő folyóiratszám, s ha ez nem jön meg a várt időre, jelzést ad, és előállítja a kiadónak küldendő sürgős levelet is.

Az online rendszer nyilvántartja a gyarapítási, megrendelési adatokat is, akár vásárolt, akár ajándék, akár csere útján beérkezett dokumentumról van szó. A teljes előfizetési adminisztrációt is immár online módszerrel végzik az INSPEC-nél, mégpedig nagy pontossággal.

/HOWES, B.: INSPEC's new acquisition system = INSPEC Matters, 6. sz. 1981. dec. p. 8./

(Roboz Péter)

AZ EURONET-DIANE rendszer adathálózati programja

Az Európai Gazdasági Közösség (EGK) Tanácsa 1971-ben határozta el egy olyan számítógéphálózat létesítését, amelyhez az EGK bármely tagországában üzemeltetett terminálokról hozzá lehet férni. A hálózat célja tudományos, műszaki és társadalmi-gazdasági adatbázisok elérése és azokban az információkeresés lehetősége. A tervezők úgy képelték, hogy a hálózathoz nemcsak EGK tagországból lehet majd csatlakozni, továbbá az EGK-n kívüli országok adatbázisainak elérése is kívánatos.