



Az „altera pars”

A Micro-ISIS lehetőségei újabb változatainak tükrében*

A FREETEXT-ről szóló cikk külön fejezetet szentel a Micro-ISIS-szel való összehasonlításnak. Az összehasonlítás nyilván a Micro-ISIS legelső, s Magyarországon leginkább elterjedt változatára, az 1.0 verzióra vonatkozik. Azóta azonban megjelent a fejlettebb változata is, sajnálatosan nem a szakmában szokásos korrekt jelöléssel, amely 1.1 lett volna. Jobb híján ezt 1987. tavaszi változatnak nevezhetjük. Több olyan fejlesztést tartalmaz, amelyet a cikk még hiányosságként jelez. Éppen az ebben és a további fejlesztési munkában való részvételre kapott szakértői meghívást a SZÁMALK osztályvezetője, Szűcs András a program fejlesztőjétől, ill. az UNESCO-tól.

E cikk megjelenésekor már várhatóan rendelkezésre áll a 2.0 verzió is, amely nagyon jelentős újdonságokat hoz, többek között a többfelhasználós lehetőséget, 1 millió rekord kezelési képességét és a kiegészítő programok írására szolgáló, a Micro-ISIS zártságát lényegében megszüntető ISIS PASCAL nyelvet. Az újabb változatokat az UNESCO-nál közvetlenül, a SZÁMALK révén pedig közvetlenül regisztrált felhasználók igényelhetik, ill. automatikusan megkapták. Ezek ismeretében adható válasz a cikkben jelzett hiányosságokra.

A Micro-ISIS zártsága

Az UNESCO továbbra sem adja a forráskódot, de az ISIS PASCAL nyelv segítségével, a belső fájlstruktúra ismeretének hiányában is készíthetők a Micro-ISIS-be szervesen beépülő kiegészítések, bővítések, fejlesztések.

* Hozzászólás ERDŐS Iván—BISZAK Sándor: A FREETEXT szöveges információkereső rendszer c. cikkéhez. (Lásd e számban p. 90—98.)

Karakterkészlet

Az új fejlesztések sem oldják meg a teljes magyar karakterkészlet kezelését. Adatbevitelkor, tároláskor, megjelenítéskor eddig is lehetett a magyar kisbetűs, hosszú ékezetes karakterek többségét, az á, é, í, ó és ű betűket használni. Valóban természetes igény, hogy minden magyar karaktert lehessen kezelni, de eddigi tapasztalataink szerint az óriási egyéb előnyök mellett ez a hiányosság eltörpül.

Kereséskor eddig sem és a jövőben sem lehet(ett) az ékezetes formát megkülönböztetni az ékezet nélkülitől. Bár eddigi gyakorlatunkban ez nem okozott több zajt a kereséskor, mint a teljes ékezetes karakterkészlettel is kivédhetetlen homonimák, valóban nem szép korlátozás.

A rendezés eddig is megvalósítható volt a magyar szabvány szerint, mégpedig az igen színvonalas csiki-csuki módszerrel. Ez lehetővé teszi, hogy külön rendezési kódsorozatot lehessen megadni magában a rekordban, amely azonban a kinyomtatáskor nem jelenik meg, tehát nem csúfítja el a rekordot. A 2.0 változatban külön rendezési kódtáblázatot is lehet készíteni.

Adatellenőrzés

Korlátozott, csak formai adatellenőrzést már az 1.0 változatban is lehetett készíteni az adatelem típusa, hossza, sémája szerint, ill. lehetett ellenőrizni a kötelező mezők és almezők hiányát. A tartalmi ellenőrzést, így az említett egyediségellenőrzést a 2.0 verzióval rendelkezésre álló ISIS PASCAL-ban megírható felhasználói programok tezik lehetővé.

Több invertált fájl

Már az alapváltozatban is ki lehetett az invertált fájl úgy alakítani, hogy — ún. prefixes invertálási módszerrel látszólag több egységre osztva az invertált fájl — bizonyos keresési fogalmakat (szerző, kiadó, nyelv, ország) elkülönítsünk benne, mintha azokat a szerzői, kiadói stb. indexben keresnénk.

Maximális rekordállomány

A 2.0 változat már 1 millió rekordot képes kezelni, ami a CD-ROM adatbázisokról való áttöltés miatt hamarosan tényleg szükségessé válhat. (A cikk ugyan nem említi, de tény, hogy az 1.0 változat még az elméleti 32 000 rekordállományt sem tudta kezelni, 4–5000 rekord körül "elrontotta" az invertált fájl. Egy saját és két külső cégnek létrehozott adatbázisnál ez komoly gondot jelentett egy-két hónapig, de az új, 1987. tavaszi verzióval megoldódott ez a probléma.)

Betekintés az invertált fájlba

A 1987. tavaszi változattól kezdve a Micro-ISIS-ben is betekinthejük az invertált fájlba a kereshető értékek közé az ún. T opció kiválasztásával. Sőt, meg is jelölhetjük azokat a keresőelemeket, amelyeket a keresőkérdés megfogalmazásakor akárunk használni, s ezeket a Micro-ISIS automatikusan átviszi a keresőkérdésbe, tehát nem kell bepötyögni, ami mindig hibát okozhat.

Szóelválasztó, ismétlési és delimiter karakterek

Mindezek már az alapverzióban megvoltak, s azóta is "könnyebbé teszik az adatok kezelését" a Micro-ISIS felhasználók számára.

Statisztikai elemzés

Mi is szükségesnek éreztük, hogy a Micro-ISIS lehetőségeit bővítsük, ezért már az 1987. tavaszi (tehát ISIS PASCAL nélküli) változathoz kifejlesztünk egy olyan programcsomagot, amely hat kiegészítő funkciót nyújt, ráépülve a Micro-ISIS-re. Ez az ISIX-PACK, amely lehetővé teszi:

- ◆ az egyes ISIS-programok (keresés, bevitel, adatbázis-definiálás stb.) használatának korlátozását felhasználói jelszók alapján;
- ◆ statisztika készítését a masterfájl rekordjairól, mezőiről (átlagos, maximális, minimális rekord-és mezőhossz, az egyes mezők előfordulási gyakorisága stb.);
- ◆ pillanatkép készítését a masterfájl rekordjainak állapotáról, az esetleges újraszervezés érdekében;
- ◆ tetszőleges rekordkészletek gyakoriságstatisztikáját önmagukban, ill. egy vagy két másik rekordrészlethez viszonyítva;
- ◆ aritmetikai műveletek végzését numerikus mezőkön, almezőkön, táblázatos output készítését;
- ◆ dBASE III fájlok konvertálását ISIS/ISO fájlkká.

A cikkben és a hozzászólásban leírtak is jelzik, hogy nincs tökéletes szoftver, mindegyik fejleszthető, bővíthető, korrigálható. Öröm, hogy az ilyen munkában egyre több hazai szakember is részt tud venni, s nem csupán passzív használóként alkalmazza a külföldön készített programokat.

Jacsó Péter
(SZÁMALK)

Felhívjuk olvasóink figyelmét, hogy a Tudományos és Műszaki Tájékoztatás 1976–1986. évfolyamainak egyes — még meglévő — számaint kérésükre térítésmentesen megküldjük, hogy hiányos évfolyamaikat kiegészíthessék.

**Budapest, Pf. 12. 1428
A TMT szerkesztősége**