

amerikai könyvtáros nagyon büszke vagyok arra, hogy ebben a nagyszabású munkában részt vehettem. Munkámmal és tudásommal nemcsak új hazámat szolgáltam, hanem hozzájárulhattam a magyar kultúra megismertetéséhez és terjesztéséhez Magyarország határain túl is.

## AZ AMERIKAI HUNGAROLÓGIAI ADATBÁZIS

BASA, PETER

Mindenekelőtt köszönetet szeretnék mondani a Magyar Könyvtárosok Egyesületének, a Magyarok Világszövetségének, és a Széchényi Könyvtárnak az értekezletre való meghívásáért.

Az eddig elhangzott előadásokban már sokan utaltak a számítógép könyvtárakban való használatára, ugyanakkor – a Kongresszusi Könyvtár előadói kivételével – senki sem tért ki részletesen ezzel kapcsolatban a saját könyvtárában szerzett tapasztalataira.

Az Egyesült Államokban ma már minden könyvtárban megtalálható a számítógép. Több specializált rendszer is létezik, ilyen például az ógörög tárgy- és szövegadatrendszer. Figyelembe véve a modern technika fejlődését és a számítógép könyvtári használatát, az Amerikai Magyar Tanárok Egyesülete 1979 őszén munkacsoportot hozott létre (Information Exchange Task Force), amelynek két fő feladatot kell ellátnia:

- hungarológiai adatbázis (Hungarian data base) létrehozása;
- Magyarország és az Egyesült Államok közötti könyvtárosi tanulmányutak szervezése és megvalósítása.

Az utóbbiról a cserélatogatási akció vezetője, Bíró Ruth fog beszámolni, én pedig a hungarológiai adatbázist szeretném ismertetni.

A vezetőség a rendszer központját a bloomingtoni Indiana Egyetemen szeretne felállítani, annál is inkább, mert ott nemrégiben a Magyar Tudományos Akadémia anyagi hozzájárulásával magyar tanszék alakult. Mivel azonban az előzetes érdeklődésre és a terv ismertetésére negatív válasz érkezett a munkacsoport a Pittsburgh-i Egyetemen próbálkozott. Az egyetem előzékeny és hozzáértő válasza alapján elindítottuk az adatbázis felállítására vonatkozó tárgyalásokat. A rendszer kidolgozásával kapcsolatos részleteket egy 1980 októberében tartandó ülésen beszéljük meg.

A hungarológiai adatbázis minden Magyarországgal és a magyarsággal kapcsolatos anyagra kiterjed (magyar nyelvű és magyar vonatkozású könyvek, folyóiratok, folyóiratcikkek, levéltári anyag, Észak-Amerikában található magyar népművészeti tárgyak stb.). A nagy közművelődési és egyetemi könyvtárak (Kongresszusi Könyvtár, OCLC stb.) anyagán kívül az összes elérhető magánkönyvtár magyar állományát is szeretnénk bevinni a rendszerbe.

A tételek bevitele (input) három módon történhet:

- mágnesszalagon;
- OCR (Optical Character Recognition – optikai karakterfelismerés) formában;
- terminál útján.

Az adatbázis létrehozása során figyelembe kellett vennünk, hogy az ne csak kutatók számára legyen elérhető, hanem mindenki számára, akit csak a magyarságra vonatkozó információk érdekelnek. Az adatbázis a rekordszerkezet tekintetében – amennyire csak lehet – a Kongresszusi Könyvtár MARC (Machine Readable Cataloging) rendszerét követi, ami megkönnyíti a jelenleg kiépülőben lévő adatfeldolgozási hálózatba való bekapcsolódást. Az adatbázis megvalósításának másik feltétele az, hogy minél olcsóbban lehessen hozzájutni; éppen ezért igyekszünk részt venni a megosztott adatfeldolgozásban (Distributed Data Processing, DDP) is. A hálózathoz való kapcsolódással a rendszert elérhetővé tennénk az Egyesült Államokon kívül is.

A legnagyobb munka az anyag minőségének ellenőrzése, amit sajnos számítógéppel nem lehet megoldani, legalább is nem 100%-osan. Vannak bizonyos használható ellenőrző formák, de a végső ellenőrző munka mégis a szakemberekre vár. Ahogy a számítógéppel kapcsolatban mondjuk: „garbage in, garbage out”, azaz amennyiben az input nem szakszerű és hiteles, úgy a visszakeresett anyag sem lehet hiteles.

Az információk visszakeresési lehetőségének megvalósítása nem jelentett különösebb problémát, hiszen ilyen programrendszerek százával állnak rendelkezésre a piacon. Eddig két rendszert választottunk ki, de választásunk nem végleges. Az egyik a Pittsburgh-i Egyetem kifejlesztett PIRATES elnevezésű software, a másik pedig az SLR (Symbolic Logic Retrieval) programrendszer. A PIRATES rendszer csak angolul „ért”, de az SLR bármilyen természetes nyelven alkalmazható. Ez utóbbi rendszer egy olyan „szótár” létrehozásán alapszik, amelyben csak az adatelemek megnevezését (field names) és azoknak a rekordban való helyét kell megadni. Az adatelemeket (szerző, cím, kiadás helye stb.) akár angolul, akár magyarul be lehet vinni a rendszerbe, hiszen a számítógép az illető adatelem helye szerint keres.

Amennyiben pl. Enikő M. Basa Petőfiről írott könyvét szeretnénk visszakeresni, úgy a következőket gépelhetjük be a számítógépbe:

A AUTHOR (vagy SZERZŐ) – ENIKŐ M. BASA (vagy csak BASA)

B TITLE (vagy CIM) – PETŐFI.

Ebben az esetben a számítógép csak egy könyvet keres ki: Enikő M. Basa Petőfijét. Ha minden, Petőfiről szóló könyvet ki szeretnénk keresni, akkor egy speciális kódot kell alkalmazni, ez pedig a CONTAINS szó rövidítéséből a CN. Tehát:

A AUTHOR CN PETŐFI.

Ebben az esetben minden könyvet megkapunk, amelynek címében Petőfi neve előfordul. Mint említettem, a számítógépnek az illető adatelem helyét is meg kell adni, pl. SZERZŐ 1–35, ami azt jelenti, hogy a szerző neve a rekord 1–35 karaktere között található.

Az 1. ábra a hungarológiai adatbázis rendszerét mutatja. A központi számítógép tartalmazza az összes anyagot, a szatellit-számítógépek pedig a helyi anyagot tárolják. Amennyiben a helyi számítógépben nem található a kért anyag, ez átkapcsol a központi számítógéphez. Természetesen a felhasználó a központi számítógépen is kezdheti a keresést, ha ez neki gazdaságosabbnak tűnik.

Nyilván nem számíthatunk arra, hogy a rendszer összes számítógépe hasonló rendszerű legyen. Éppen ezért kidolgoztunk egy olyan alrendszert (sub-system), amely

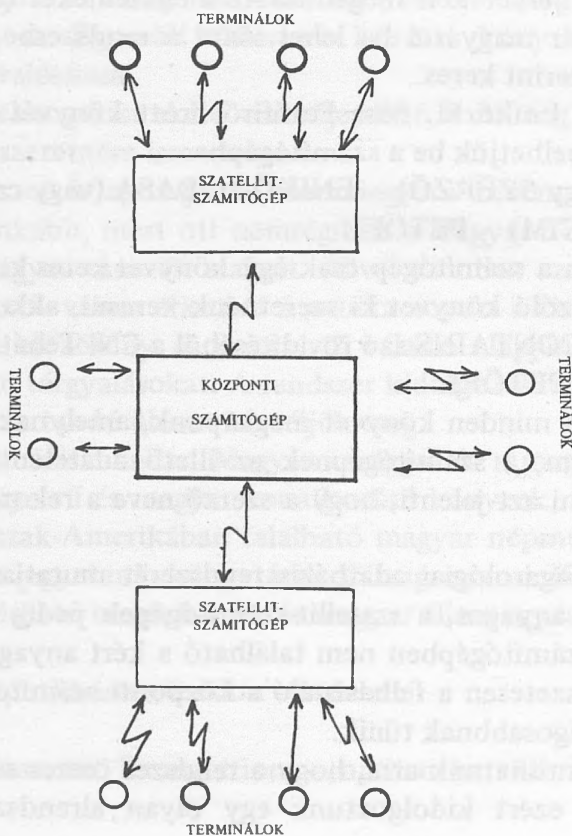
a különböző számítógépek között „tolmácsol” (2. ábra). Ez az alrendszer kezeli az anyagot, ha például egy 8-as rendszerű számítógépről egy 16-os rendszerűre kell váltani, mint például egy DEC típusról egy IBM-re. Mivel ilyen „fordító” rendszerek készen kaphatók, kevés módosítással – és így aránylag olcsón – használhatók.

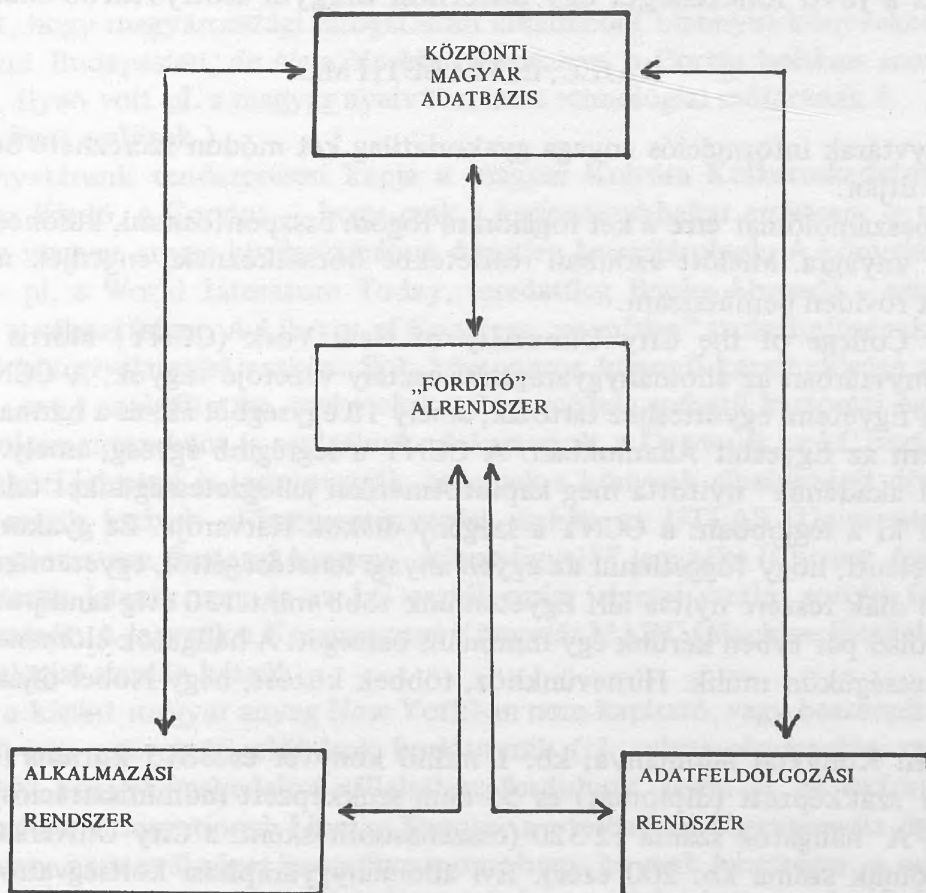
A rendszer 1981 nyarán kezdi meg üzemszerű működését; ekkor köteget (batch) üzemmódú keresésre lesz lehetőség. Ez időre az adatbázis előreláthatóan már tartalmazni fogja a legnagyobb könyvtáraktól kapott mágnesszalagokat.

A munka teljes sikeréhez az Amerikai Magyar Tanárok Egyesületének munkacsoportja szeretné megkérni az ezen értekezleten résztvevőket, a Magyarok Világszövetségét, a Magyar Könyvtárosok Egyesületét, a Széchényi Könyvtárat és mindeki mást, hogy aki teheti, ebben a munkában legyen segítségünkre.

Miben állna ez a segítség? Lényegében cseréről és javaslatokról van szó. Szeretnénk valamilyen formában megkapni az adatbázisunkba bevihető magyar anyagot, de cserébe mi is adunk egy modern rendszert, amely némi változásokkal Magyarországon is használható az „R” számítógépeken, és átadjuk a rendszerben tárolt anyagot is bármilyen helyi feldolgozásra. A sajtótól már többen kérdezték, hogy mit kérünk Magyarországtól. Pár szóval ezelőtt, igaz, „kértem”, de szeretném hangsúlyozni, hogy az Information Exchange Task Force nem kérésekkel jött, hanem azzal a szándékkal, hogy információt cseréljünk.

1. ábra





## IRODALOM

1. BURCH, J. G. Jr.,—STRATER, F. R.—GRUDNITSKI, G.: *Information Systems: Theory and Practice*. New York, J. Wiley and Sons, 1979.
2. CHENHALL, Robert G.: *Nomenclature for Musieum Cataloguing; a System for Classifying Man-made Objects*. Nashville, American Association for State and Local History, 1978.
3. HOCKEY, S.: *A Guide for Computer Applications in the Humanities*. Baltimore. Johns Hopkins University Press, 1980.
4. KAIZAN, M. Jr.: *An Introduction to Distributed Data Processing*. New York, Petrocelli Book, 1978.
5. OAKMAN, R. L.: *Computer Methods for Library Research*. Columbia, U. of South Carolina Press, 1980.