

A Magyar Nemzeti Bibliográfia (MNB) számítógépes rendszere

(Nyáryné Grófcsik Erika, Sipos Márta)

A projekt háttere, célja, életrajza

A UBC (Universal Bibliographic Control = Egyetemes Bibliográfiai Számbavétel) eszméje már a XIX. század végén felmerült, hivatalos, az IFLA és az UNESCO által is támogatott programként 1974-ben fogalmazódott meg. Célja, hogy a világon bárhol megjelent dokumentum bibliográfiai adatai a nemzeti bibliográfiák világméretű integrált hálózatán keresztül bárhol hozzáférhetőek legyenek. A nemzeti bibliográfiák felelőssége egy ország területén (esetleg adott nyelven, vagy egy ország területén kívül adott etnikumra vonatkozóan) megjelent dokumentumok elsődleges, teljességre törekvő és hiteles regisztrálása, ami – általában - nemzeti könyvtári feladat és döntően a kötelezpéldány szolgáltatásra épül. A UBC program a nemzeti bibliográfiák számára számos tartalmi és formai irányelvet határozott meg: a dokumentumtípusok közül a könyvek és könyvjellegű dokumentumok, új indulású periodikumok kurrens bibliográfiájának nyomtatott formában történő közzétételét kötelezően előírta, a bibliográfiai leírásokat egységes, szabványos formában (ISBD), szabványos központozással követelte meg, elvárta a dokumentumok azonosítását megkönnyítő szabványos könyv- és folyóirat-azonosítók (ISBN, ISSN) alkalmazását. A hetvenes évek elején az angolszász országok nagyobb könyvtáraiban általánossá vált a számítógép alkalmazása és előbb a Library of Congress, majd a British Library megindította az általa előállított kurrens nemzeti bibliográfiai adatainak elektronikus formában, mágnesszalagon történő szolgáltatását a számítógépes rendszert működtető könyvtárak számára. Ekkor született meg – az ISO 2709 implementációjaként - a MARC (Machine Readable Cataloging) rekordstruktúra, ami a bibliográfiai adatok, ezen belül a bibliográfiai leírás, a besorolási adatok és a példányadatok kódolásának és rekordon belüli elhelyezésének, ill. ábrázolásának módját szabályozza. Bár a MARC struktúra az angolszász, főleg a Library of Congress (1970 körüli) feldolgozó rendszerének sajátosságait tükrözi, de facto szabvánnyá, a UBC program keretében 1977-ben publikált, nemzetközi adatcserére kidolgozott UNIMARC formátum és számos nemzeti MARC-változat - többek között a 2002-ben magyar szabvánnyá minősített HUNMARC - kiindulási pontjává vált és ma is széles körben használt.

Az Országos Széchényi Könyvtár 1973-ban kereste meg az SZKI-t azzal az ajánlattal, hogy együttműködésben hozza létre a nagyműltű „Magyar Nemzeti Bibliográfia. Könyvek Bibliográfiája” című, a kurrens magyarországi könyvkiadást regisztráló, félhavonta megjelenő folyóiratot, annak éves kumulációjait és ciklusbibliográfiáit, a Magyar Könyvészeteket előállító számítógépes rendszert oly módon, hogy az minden szempontból feleljen meg a UBC kívánalmaknak, és egyben a magyar nyelv sajátosságainak, valamint bizonyos magyar könyvtári előírásoknak, tradícióknak is. A feladatot az akkor még újak számító „egyszeri feldolgozás, többszöri felhasználás” alapelvének következetes érvényesítésével, egy bibliográfiai adatbázis felépítésével kívánták végrehajtani.

A bibliográfiai kiadványok ebben az időben teljesen manuálisan, egy IBM gömbfejes írógéppel, házon belüli sokszorosítással készültek, és hasonlóan manuálisan történt minden egyéb könyvtári feldolgozás is. Előtanulmányok, munkaszervezési változtatások, új feldolgozási szabályok, szabályzatok, sőt szabványok kidolgozása, speciális input-output berendezések felkutatása, ill. beszerzése volt szükséges ahhoz, hogy a programfejlesztés megindulhasson. A rendszer tesztelése az 1976-os állomány feldolgozásával kezdődött, az üzemszerű működés 1978-ban, a 33. évfolyammal indult.

Az akkori viszonyok között a rendszer működtetése elképzelhetetlen lett volna az SZKI-n belüli és kívüli partnerek szoros és jó együttműködése nélkül. A bibliográfiai adatokat a könyvtárosok adatlapra

írták, ezt követte az adatrögzítés, ami az OSZK-ban felszerelt, a Nyomdaipari Fényszedő Üzemben is használt Monotype Multicode 913 adatrögzítő berendezésén történt. Ez egy nagy jelkészlet bevitelére alkalmas, kb. 90 (!) billentyűvel felszerelt, megjelenítő nélküli 8 csatornás lyukszalag lyukasztó gép volt. A lyukszalagokat a KERSZI RC600-as gépén mágnesszalagra másolták, amely aztán hetente egyszer az SZKI Számítógép Laboratóriumának (SZL) Siemens 7755-ös számítógépén, BS2000 batch üzemmódban, az SZKI Software Alkalmazások Laboratóriumának (SAL) munkatársai által fejlesztett programokkal került feldolgozásra. A feldolgozás az input adatok ellenőrzéséből, a bibliográfiai adatbázis naprakésszé tételéből és az ellenőrző listák előállításából állt. Ez utóbbi mágnesszalagra készült, mivel a nagy jelkészlet miatt csak az OSZK által beszerzett, speciális, 163 jel nyomtatására alkalmas IBM/ALA (American Library Association = Amerikai Könyvtáros Egyesület) nyomtatólánccal lehetett az ÉGSZI IBM 370/125-ös berendezésén kinyomtatni. Havonta kétszer az input adatok feldolgozását egy második fázis, a bibliográfiai kiadvány előállítása követte, amely teljesen automatikusan, egy algoritmus alapján kiválogatta a bibliográfiai adatbázisból a megjelentethető rekordokat, azokból az adott kiadvány (paraméterekkel definiált) szerkesztési szabályainak megfelelően felépítette a kiadványt és kiírta mágnesszalagra a mindenkor nyomtató/megjelenítő berendezés kívánalmi szerint konvertáltan és formátáló parancsokkal ellátva. Kezdetben a mágnesszalagról az ÉGSZI-ben, a speciális ALA sornyomtató-lánccal készült a füzet, egy-két év után azonban a Nyomdaipari Fényszedő Üzem már tudott mágnesszalagot fogadni. Számítógépes rendszerükkel feldolgozták, majd a Monophoto600-as berendezéssel megjelenítették („levilágították”) és papírra fotózták annak tartalmát, a tőpéldányt pedig az OSZK nyomdájában sokszorosították. Évente egyszer ugyanezzel az eljárással, de némileg eltérő kritériumok alapján végrehajtott automatikus válogatást és szerkesztést követően a teljes év anyagából egy könyv formájú (többkötetes) bibliográfia is készült.

Az MNB rendszer 1991-ig, tizennégy éven keresztül – folyamatosan – üzemelt. Ezalatt számos új szolgáltatással bővült: a bibliográfiai rekordokból az ALA nyomtatólánccal katalógus cédulák készültek a Könyvtárellátón keresztül az ország (több mint háromezer) közművelődési könyvtára, valamint az OSZK nyomtatott katalógusai számára, az MNB adatbázisa szolgáltatott adatokat az Új könyvek c. állománygyarapítási tanácsadó folyóirathoz, a teljes 1976-tól 1991-ig terjedő időszakról pedig egy hatalmas ciklus-bibliográfia készült microfiche-re a Volán Elektronika berendezésén. Jelentősen javult a rendszer „komfort-fokozata” is: 1982-ben a Monotype lyukasztó berendezést az SZKI Hardware Laboratóriumában (HL) fejlesztett és az OSZK igényekhez adaptált TS-51 intelligens terminál (2, majd 4 munkaállomással) váltotta ki, amelyről a rögzített adatokat adatátviteli vonalon keresztül közvetlenül az SZKI/SZL számítóközpontjába lehetett küldeni. A TS-51 terminálhoz bővebb jelkészlet kinyomtatására alkalmassá tett DARO és FACIT nyomtatókat illesztettek, így az ellenőrző listákat a TS-51 munkaállomásokra vissza lehetett küldeni és helyben kinyomtatni.

1992-ben az OSZK saját IBM számítógépet és standard könyvtári rendszert (DOBIS/LIBIS) kapott, amelyet az ezredfordulón a jelenleg is működő, nagy teljesítményű szervereken futó, ORACLE adatbáziskezelőre épülő és a web-felületű LibriVision keresővel kiegészített Amicus rendszer váltott fel. Az 1976-tól 1991-ig feldolgozott és archivált állományt az OSZK – az MNB rendszer hajdani fejlesztőinek segítségével – az új rendszerbe migrálta, ott annak adatai ma is hozzáférhetőek. Ugyanígy kifejlesztésre került egy új, a régi MNB rendszer tapasztalataira, módszereire épülő, új nemzeti bibliográfiai kiadvány rendszer, amely ma már csak web felületen keresztül olvasható, de kiterjed további dokumentum típusokra (folyóiratok, térképek, kották, hangfelvételek), és biztosítja az egyedi és ingyenes hozzáférést a bibliográfiai rekordok MARC (HUNMARC) formátumához is.

Az MNB rendszer számítástechnikai újdonságai, kihívásai

Korunkban, a XXI. század elején divat a Gutenberg galaxis eltűnéséről beszélni, a könyvtárakat tudástáraknak vagy tudásbázisoknak nevezni, a katalógusok és bibliográfiák helyett közvetlenül a kiadványok digitálisan elérhető tartalmát böngészni és a legkisebb személyi számítógépen is nyomdai minőségű dokumentumokat előállítani. Mindez fantasztikus fejlődés a fent vázolt kezdetekhez képest, érdemes azonban belegondolni, hogy az akkori HW/SW technológiai fejlettségi szinten milyen problémákkal kellett szembenézni. Íme néhány, az akkor nem triviálisan megoldható feladatok közül:

- **Változó összetételű és hosszúságú adatstruktúrák (rekordok) egymással logikai kapcsolatban**

A bibliográfiai rekordok adatainak azonosítására a MARC adatcsoport- és adatazonosítók és az ezeket tovább finomító indikátorok, valamint az adattartalomba (stringekbe) ágyazott almező azonosítók szolgálnak, amelyek együttesen rendkívül sokféle rekordfelépítést eredményeznek. Az adattartalom alapján a bibliográfiai rekordok között különféle logikai kapcsolatok létesíthetők. További sajátosság, hogy a fenti rendszerben azonosítható adatelemek többsége egy karaktorsor (string), amelynek hossza igen tág határok között változik. Mindezek hatékony kezelésére a hetvenes évek második felében nem álltak rendelkezésre megfelelő adatbázis-kezelő rendszerek és szabványos magas szintű programozási nyelvek. Mind az adatbázis kezelő rendszert, mind a feldolgozásokat végző kód jelentős részét – a mintaként szolgáló nagy könyvtárakhoz (pl. British Library, Library of Congress) hasonlóan – ASSEMBLER nyelven kellett programozni. A könnyebb karbantarthatóság / továbbfejleszhetőség érdekében külön input-output rendszert fejlesztettünk ki, a programok pedig táblázat-vezérelt módon működtek.

- Rendkívül nagy jelkészlet

A magyar nyelv ékezetes betűinek korrekt megjelenítése a magyar nemzeti bibliográfiákban alapvető és vitathatatlan követelmény. A UBC irányelvek és a szabványos bibliográfia leírás (ISBD) azonban azt is megköveteli, hogy a Magyarországon idegen nyelven kiadott művek adatai vagy a fordításban megjelent művek eredeti címei az eredeti nyelv írásmódjának megfelelő betűkkel és ékezetekkel – nem-latin ábécé esetén ISO transliterálással – jelenjenek meg. Ezt támogatandó dolgozták ki a hetvenes években az ISO 5426 jelzetű, bibliográfiai használatra szánt kódtáblát, amely nem tartalmaz ékezetes betűket, de minden latin ábécében használt ékezethez – szám szerint 31 – önálló kódpozíciót rendel. Az ékezetes betűket két karakterrel – ékezet és alapjel – lehetett ábrázolni, így ezek valamennyi kombinációja elvileg megengedett volt. Az MNB rendszerben ez a kódolási mód került alkalmazásra. A legnagyobb problémát az adatok bevitele, még inkább azok megjelenítése jelentette. A hetvenes években a nyomtatók zöme csak 64 karaktert (számok, az angol ábécé nagybetűi, néhány írásjel és szimbólum) ismert. Megoldást a már említett, Magyarországon unikumnak számító IBM/ALA nyomtatólánc jelentett, amely önálló jelként tartalmazta az ékezeteket. Minden sort két sorként – alapjel-sor és ékezet-sor – kellett felépíteni és ezeket soremelés nélkül egymásra nyomtatni. A Nyomdaipari Fényszedő Üzem - egy korábbi bér munkában végzett lexikonszedésnek köszönhetően – rendelkezett egy olyan levilágítható betűkészlettel, amely tartalmazta az európai nyelvekben gyakrabban előforduló ékezetes betűket. Mivel ez a jelkészlet csak néhány betűtípusban és méretben állt rendelkezésre, az MNB kiadványok tipográfiáját ennek megfelelően kellett megtervezni, illetve megvalósítani.

- Besorolás (rendezés)

Az MNB kiadványok a bibliográfiai leírásokat szakrendben, ezen belül betűrendben tartalmazzák és különféle, a bibliográfiai leírások megtalálását könnyítő, a könyvtári besorolási szabályok szerint összeállított betűrendes mutatókkal egészülnek ki. Ezek számítógépes rendezése számos olyan (ma is élő) problémát vetett fel, amely csak különleges eljárások kidolgozásával volt megoldható. Ilyen az ékezetek és központosági jelek figyelmen kívül hagyása, a címeket kezdő határozott és határozatlan, magyar- és idegen nyelvű névelők átugrása, a szövegbe ágyazott számértékek helyes rendezése, a római számok felismerése és arab számokként történő rendezése, stb.

A mai mércével is igen komplex MNB rendszer hosszú, sikeres működésének titka és legnagyobb erénye az üzembiztosság volt: minden hónap 1-én és 15-én megjelent az MNB füzet. Köszönhető ez a megrendelő OSZK és a szolgáltató SZKI közt mind vezetői, mind munkatársi szinten kiváló együttműködésnek és az SZKI környezetnek, amely minden tevékenységben minőséget követelt és minőséget támogatott. Örömmel gondolunk vissza arra, hogy ez a szinergia más irányban is működött: amikor a RECOGNITA egyre sikeresebb lett és bővíteni akarta jelfelismerő képességét az idegen nyelvű szövegekben, az MNB fejlesztők tudtak teljes körű és pontos információt, megfelelő példákat adni a latin ábécét használó nyelvek speciális és ékezetes betűiről.