

A NEMZETKÖZILEG HOZZÁFÉRHETŐ ADATBÁZISOK REKORDJAINAK SZERKESZTÉSE INTELLIGENS TERMINÁLOKON*

Alan Hopkinson

British Library, London

1. Az adatszere-tevékenység fejlődése

Az utóbbi 15 évben számos erőfeszítést tettek annak érdekében, hogy meghatározzák a bibliográfiai adatok géppel olvasható formában való cseréjének egyes folyamatait. Bibliográfiai adatok alatt bibliográfiai rekordokat, könyvekre, folyóiratokra, filmekre, hangfelvételekre, folyóiratcikkekre vonatkozó hivatkozásokat értünk, melyeket szerzőjük vagy szerkesztőjük neve, vagy témájuk, tehát címük, kulcsszavak, osztályozó rendszerek jelzetei, illetve szabályozott nyelvi rendszerek pl. tezauszok és tárgykör csoportok deskriptorai szerint indexelünk. Figyelemre méltó, hogy milyen sok munkát fordítottak e rekordok kezelési folyamatainak kifejlesztésére, melyek azonban sohasem, vagy csak igen ritkán jelentik a keresés végpontját, hiszen a keresés célja minden esetben a bibliográfiai rekordban hivatkozott dokumentum. Akár kutatóról van szó, akinek a rákellenes hatóanyagokról 1983-ban megjelent folyóiratcikkekre van szüksége, akár arról a személyről, aki a közművelődési könyvtár katalógusában egy regényt vagy egy olyan könyvet keres, melyből megtudhatja, hogyan újítsa fel háza elektromos vezetékeit, *sohasem a bibliográfiai rekord az, amit keres, hanem az eredeti közlemény*. Másokra hagyom annak a boldogtalan időszaknak megjövendölését, amikor már nem lesz sem könyvünk, sem újságunk, csak fogjuk az ágyunk melletti telefonkészüléket és képernyőre hívjuk kedvenc regényünket.

Biztos vagyok benne, hogy még sok idő fog eltelni addig, amíg az elektronikus sajtó szükségtelessé teszi a nyomtatott irodalmat, s ha így is fog történni, még mindig szükségünk lesz nyomtatott indexekre, ha nem akarunk elmerülni a különböző tudományterületeken felhalmozott ismeretek tömegében.

A bibliográfiai rekord tehát jelen van a tömeges adattárolásra alkalmas eszközök korában is, és e cikkben

éppen azt szeretném bemutatni, hogy a modern technika hogyan járulhat hozzá az egyetemes bibliográfiai számbavétel céljainak megvalósulásához (bővebbet lásd pl. Clarke-nál [1] azáltal, hogy mind szélesebb körben teszi hozzáférhetővé a világ bibliográfiai szervezetei által létrehozott bibliográfiai rekordokat.

2. A bibliográfiai adatszere jelenlegi helyzete

Napjainkban sok, bibliográfiai adatszerevel foglalkozó rendszer létezik. Az első nagyszabású erőfeszítéseket az Egyesült Államokban, a Library of Congress-nél tették. 1966-ban kidolgozták a MARC csereformátumot [2], majd kísérleti alkalmazás után, 1968-ban tökéletesítették, együttműködve a British National Bibliography Ltd.-el, amely az angol nemzeti bibliográfiát előállító nem nyereségérdekelt magánvállalkozás volt. Az ily módon létrejött tulajdonképpeni első igazán nemzetközi csereformátum segítségével lehetővé vált, hogy a BNB, majd később a British Library fogadhassa a Library of Congress rekordjait és beépítse azokat a más intézmények számára készített katalógusaiba.

H. D. Avram, aki támogatta a MARC kidolgozását a Library of Congress-nél [3], az UNIMARC dokumentum [4] előszavában a csereformátumoknak az alábbi elemeit határozta meg:

- *a rekord szerkezete*, amely az információ fizikai megjelenítése a géppel olvasható hordozón;
- *a rekord adatelem-azonosítói*, melyek az adatelemek azonosítására, illetve az adatelemekre vonatkozó kiegészítő információk adására szolgálnak;
- *a rekord tartalma*, amely maga az adat, pl. a szerző neve, a dokumentum címe stb.

Az alábbiakban vegyük sorba az egyes elemeket, és vizsgáljuk meg, hogy melyik milyen mértékig szabványosított.

* A DATABASE '83 c. nemzetközi konferencián (Budapest, 1983. június 6-8.) elhangzott előadás.

2.1 Rekordszerkezet

A rekordszerkezetet illetően szintén általános volt az egyetértés, amióta 1973-ban az ISO elfogadta az ISO 2709 szerinti rekordszerkezetet [5] a bibliográfiai adatcsere céljaira. Néhány országban ezt a rekordszerkezetet már korábban nemzeti szabványként fogadták el.

2.2 Adatelem-azonosítók

A rekordok adatelem-azonosítói (közismertebben a hívójelek) vonatkozásában már nem volt olyan általános az egyetértés, és ma is többféle adatelem-azonosító rendszer létezik. A nemzetközi rendszerek általában MARC-ra vagy a Reference Manual-re alapozottak, és az adatelem-azonosítókat az egyes felhasználói csoportok szükségleteinek megfelelően dolgozzák ki.

2.2.1 MARC

A MARC a világ számos nemzeti könyvtárában használt hívójelek rendszerének generikus neve. Ugyanezt a rendszert használják a nemzeti könyvtárak „ügyfelei” is, tehát az egyes országok könyvtárai, akik a nemzeti adatbázis rekordjait felhasználják. A British Library-ben és a Library of Congress-ben használt rendszer kissé elfogult a könyvtári katalógusok, illetve a nemzeti bibliográfia számára készülő rekordok javára. A hívójelek sorrendje követi a katalógus tételeinek összeállításához szükséges adatok sorrendjét, és érvényesül a rendszerben a főtételek és melléktételek gondolata is. Több országnak van saját MARC formátuma, amelyek egymással többé-kevésbé kompatibilisek. A KGST országokban használatos nemzetközi csereformátum, a MEKOF is a MARC-család tagja. A British Library átveszi a Library of Congress-ben készült rekordokat, és viszonylag csekély nehézségek árán US MARC-ból UK MARC formátumra alakítja át őket. A sok kétoldalú konvertálás problémájának elkerülésére az IFLA katalogizálási és gépesítési szekciói létrehozták az UNIMARC néven ismert nemzetközi MARC formátumot [6]. Ezzel egyidejűleg átrendezték a hívójelek sorrendjét és ezáltal az UNIMARC inkább vált alkalmassá általános célú felhasználásra, mint az egyes MARC formátumok, amelyekkel szükségszerűen kompatibilisnek kell lennie.

2.2.2 A Reference Manual

A Reference Manual-t eredetileg az UNISIST/ICSU-AB bibliográfiai leírásokkal foglalkozó munkacsoportja (Working Group on Bibliographic Descriptions) dolgozta ki. Ebben a munkacsoportban olyan szakértők vettek részt, akik magukkal hozták a mindazokban az intézményekben rendelkezésre álló tapasztalatokat, amelyek már kifejlesztettek gépesített információfeldolgozó rendszereket, különösen a referáló és indexelő szolgáltatások területén. A Reference Manual [7] második kiadása már nem kizárólag egy speciális felhasználói csoport szükség-

leteit kívánja szolgálni és nem is csak egy bizonyos fajta output előállításra alkalmas. Egyaránt alkalmas például egy folyóiratcikk és egy közönséges monográfia leírására. A szabványos rekordszerkezet (ISO 2709), valamint a rekordok szabványos összekapcsolási lehetőségének hiánya azonban hátrányosan befolyásolta fejlődését, így nehezen alkalmazható összetett rekordszerkezetek, például az egy folyóirat egy számában közölt összes cikk esetében. A szerkezet megköveteli, hogy a rekordokat a négy bibliográfiai szint (analitikus, monografikus, időszaki és gyűjteményes) szerint kategorizálják, és bár a rekordok többsége beleillik ebbe a szerkezetbe, néhány azonban nem. A leggyakoribb adatelemek esetében, mint például a szerző vagy a cím, minden szinten ismétli az adatelem-azonosítót, más adatelemek esetében azonban, mint amilyen például a disszertációk jegyzetben közölt kiadója, nem ad lehetőséget a szint kifejezésére. Mindezek ellenére a Reference Manual kiválóan alkalmas arra a célra, amire eredetileg létrehozták, nevezetesen a folyóiratcikkek, könyvek (monográfiák) és időszaki kiadványok (ebben a sorrendben) rekordjainak előállítására a referáló és indexelő szolgáltatások adatbázisai közötti információcsere céljaira.

2.2.3 Az Unesco Közös Adatcsere-Formátuma (CCF)

A hetvenes évek közepén az Unesco felismerte a két formátum párhuzamos létezése által okozott problémákat, ezért támogatta egy nemzetközi szimpózium összehívását a bibliográfiai csereformátumok kérdésében (International Symposium on Bibliographic Exchange Formats, Taormina, 1978. [8]). A szimpózium alkalmával sikerült összehozni az információs közösség különböző szektoraiban dolgozó szakembereket, és véget vetni annak az irányzatnak, amely a könyvtárak, valamint a referáló és indexelő szolgáltatások közötti polarizációhoz vezetett. A szimpózium ajánlásai [9] alapján az Unesco létrehozott egy ad hoc munkacsoportot egy közös adatcsere-formátum kidolgozására (Ad Hoc Group on the Establishment of a Common Communication Format). Ez a formátum időközben elkészült, és rövid időn belül publikálják. A munkacsoport hamar felismerte, hogy az információs közösség különböző csoportjainak igényei jobbra azonosak. A formátum alapja a minden rekordban kötelező adatelemek összessége, tehát azoké az adatelemeké, amelyek feltétlenül szükségesek a hivatkozott dokumentum azonosításához, és az ezekhez járuló további elemek egészítik ki a formátumot. A CCF-ben alkalmazható igen szellemes rekord-kapcsolási módszer lehetővé teszi a más, pl. MARC, MEKOF vagy Reference Manual formátum szerinti rekordok Közös Adatcsere-Formátumra való átalakítását.

2.2.4 Egyéb nemzetközi csereformátumok

Itt elsősorban az AGRIS és az INIS formátumot kell megemlíteni, amelyek különféle sajátosságaikat a Refe-

rence Manual-ból, illetve a MARC-ból köcsönzték. Ezek a formátumok a mezőgazdasági, illetve az atomenergiával kapcsolatos bibliográfiai adatok nemzetközi cseréjére szolgálnak. Bár az e formátumokban közölt adatok inputja decentralizált, a formátum alkalmazása központilag ellenőrzött, különösen a rekordok tartalmát érintő katalogizálási szabályok vonatkozásában.

2.2.5 Következtetések

Amint látható, több különböző rendszer létezik a bibliográfiai adatok cseréjének céljaira. Elméletileg a több formátum párhuzamos létezése költséges mulatság, akkor is, ha ma már könnyűszerrel írhatók olyan programok, amelyek segítségével az adatok az egyik formátumról a másikra konvertálhatók. A valóság mégis az, hogy mivel az ilyen programok írása idő- és munkáigényes, a sokféle formátum léte inkább akadályozza a bibliográfiai adatszert.

2.3 A rekord tartalma

Először az UNIMARC példáján tekintsük át, hogy mennyire szabványosított a rekordok tartalma. Az UNIMARC, mint már említettük, tulajdonképpen semleges formátumnak készült, amely igen sokféle dokumentum rekordjainak elkészítésére alkalmas, bár jelenleg speciális mezők csak a könyvek, időszaki kiadványok és térképek feldolgozására állnak rendelkezésre. Az adatok, hasonlóképpen a többi nemzetközi formátumhoz, oly módon kezelhetők, hogy azokból különféle típusú output legyen összeállítható. Az UNIMARC kidolgozóinak célja egy olyan szabályrendszer létrehozása volt, amelynek segítségével az e szerint kódolt, géppel olvasható formában tárolt adatokat könnyen át lehet alakítani a külső forrásokból rekordokat átvevő intézmények által használt egyéb formátumokra. Ez lehetővé teszi, hogy az, aki a rekordokat átveszi, tudja, hogy mire számíthat, és azonosíthassa az egyes adatelemeket. A cserélt rekordok életének van azonban egy másik fázisa is. Az átvett rekordot be kell építeni a fogadó intézmény adatbázisába. Azáltal, hogy az adatelemek azonosíthatók, ez a folyamat minden bizonnyal elvégezhető a legágabb értelemben. Ha a fogadó adatbázis online üzemmódú és invertált fájlja, vagy szerző, cím, illetve nemzetközi azonosító számozások szerinti indexe van, akkor ezek az adatok a rekordokban rendelkezésre állnak. Ha azonban a rekordokat nyomtatott vagy mikroformátumú katalógusba vagy bibliográfiába kívánják beépíteni, a különféle forrásokból vett rekordokról egy szempillantás alatt kiderül, hogy még véletlenül sem annyira kompatibilisek, mint szeretnénk. Az UNIMARC ez esetben is példának tekinthető. Ezt a formátumot a könyvtári közösség használja, használatát speciálisan a nemzeti könyvtárak támogatják. Ez ideig számos ország, így például Ausztrália, Kanada, Japán, Magyarország, Dél-Afrika, Nagy-

Britannia és az Egyesült Államok nemzeti könyvtára fogadta el az UNIMARC-ot csereformátumként azzal, hogy használatát az 1980-as évek elején megkezdi [10].

Mindazok, amire a következőkben ki szeretnék térni, a bibliográfiai intézmények között cserélt rekordok esetében igazak. A nemzeti könyvtári közösségen belül az egyes könyvtárak rekordjaikat általában azonos célra használják, pl. nyomtatott bibliográfiák vagy online információkeresés céljaira. Bár e rekordok azonos célra készülnek, gyakran azonban különböző katalogizálási szabályok alapján, és ez az a pont, ahol a rekordok tartalmával kapcsolatos problémák gyökereznek. A problémák adódhatnak egy szükségesnek tartott adatelem hiányából (az alkalmazott katalogizálási szabály nem tekinti fontosnak a kérdéses adatelemet), valamint az adatok közlésmódjának különbözőségeiből. Egy még nehezebben megoldható probléma az, hogy ha az adott formátumban szerepel ugyan a kérdéses adatelem, hiába alkalmas például katalóguskészítés céljaira, még nem biztos, hogy a visszakeresés szempontjából is megfelelően van kódolva. Annak illusztrálására, hogy mire is gondolok, vegyük példának a *hivatalos kiadványok azonosító számozása* elnevezésű adatelem esetét.

2.3.1 Hivatalos kiadványok azonosító számozása

Minden rendszer feltünteti a bibliográfiai tétel címét, sőt újabban az ISBN-t és ISSN-t is, mely utóbbiak a manuális rendszerekkel ellentétben a gépi rendszereknél már teljesen szokványosak. E cikkben a gépi rendszerekkel foglalkozunk. A szerző(k) neve a másik olyan adatelem-csoport, amely szinte mindig szerepel a rekordban. Ezen a néhány adatelemen kívül azonban még nagyon sok más adatelem is létezik, melyeket az egyik intézmény figyelembe vesz, a másik viszont nem. Erre példa a hivatalos kiadványok azonosító számozása is. A referáló és indexelő szolgáltatások általában szívesen használják ezeket a számokat, mivel adatbázisaik használói gyakran e számok alapján azonosítják a dokumentumokat, emellett az e fajta kiadványoknak gyakran nincs is nemzetközi azonosító számozásuk. A legtöbb könyvtári katalogizálási szabályzat előírja ugyan e számok feltüntetését, de mivel e szabályok egyúttal a rekord formai elrendezését is meghatározzák, a géppel olvasható rekordokban e számoknak nincsen azonosító funkciója, hanem rendszerint a sorozati adatok helyén, vagy megjegyzésként szerepelnek, hasonlóan a sorozatcímbe (ebben az esetben a cím gyakran egy kód, pl. Cmnd) és a dokumentum sorozaton belüli számából álló sorozati adathoz. Sok esetben a katalóguskészítő vagy a hivatalos kiadványokat ismerő könyvtáros rögtön felismeri, hogy ez a szám a hivatalos kiadványok azonosító számozása, de mivel a rekord külső forrásból származik és az adatbázisba való beépítés automatikus, minden, különösen pedig a rendszeres emberi beavatkozás időpocsékolásnak és gazdaságtalannak tűnik. Ezért ha a bibliográfiai

fájlt a hivatalos kiadványok azonosító számozása szerint kívánjuk elrendezni, akkor ezt csak olyan forrásból vett rekordokkal tehetjük meg, ahol ezt az adatelemet elkülönítve tüntetik fel. Az UK MARC-ban van jegyzetben közölt szám feltüntetésére szolgáló mező. Ez két almezőre oszlik, melyek közül a második tartalmazza a számot, az első pedig egy leíró jellegű közlés számára van fenntartva, ami lehet pl. „hivatalos kiadványok azonosító számozása”. Az utóbbi mezőt azonban nagyon kevésbé használják, mint azt a BLAISE adatbázisban végzett keresés során tapasztalhattam.

2.3.2 Tárgy szerinti hozzáférés

Ugyanez, vagy még rosszabb a helyzet a tárgy szerinti hozzáférést biztosító adatmezők esetében. Tárgyköri csoportokra (tárgyszavakra) vonatkozó nemzetközi szabvány nincsen. Vannak ugyan teaurusz-készítésre vonatkozó szabványok, nincs azonban egyetlen, általánosan használt teaurusz sem. Az ETO nemzetközi szabvány rangjára emelt osztályozó rendszer, de viszonylag kevés rekordot osztályoznak ETO szerint, és ha teszik is, nincsen garancia az egységes és következetes alkalmazásra. Így fordulhat elő, hogy bár az adatelem rendelkezésre áll az UNIMARC, a Reference Manual vagy a CCF szerinti rekordban, mégis gyakran hiányzik az e formátumokra konvertált rekordokból, még akkor is, ha feltüntetésük lehetséges lett volna.

2.3.3 Eltérő formájú rekord-tartalmak

Tekintsük most át azokat az adatelemeket, amelyek következetesen szerepelnek a rekordokban. Ezek elsősorban a szerző, a cím, és az azonosító számozások, bár ez utóbbiak nem szerepelnek sem a régebbi dokumentumokon, sem az olyan országokból származókon, ahol az ISBN használata még nem elfogadott, sem pedig azokon a dokumentumokon, amelyeket nem a szokásos kiadói csatornákon keresztül publikálnak. Ha az adatbázisokban egy dokumentum címét keressük, azt tapasztaljuk, hogy az is lehet különféle. Angol nyelvű adatbázisban találhatunk angol nyelvű címet ezzel a megjegyzéssel: „orosz nyelvű”. Világos, hogy a rekord készítője nem az eredeti címet használta, hanem csupán annak angol fordítását. Ha a szerzőket tekintjük át, azt látjuk, hogy ők is különféleképpen szerepelnek. Vegyünk példaként egy mindenki által ismert szerzőt, William Shakespeare-t, akinek neve lehet Shakespeare, William; vagy Shakespeare, W. Az ő neve alatt oly sok mű szerepel, hogy egy nyomtatott katalógusban, amely különféle forrásokból vett rekordokat tartalmaz, amelyekben a név különbözőképpen szerepel, több oldal is lehet az egyik és a másik formában közölt név alatt felsorolt rekordok között. További bonyodalmak származhatnak abból is, ha – amint ez könyvtári katalógusokban szokásos – a szerző neve mellett feltüntetik születésének évét, hogy ily módon megkülönböztethessék más, azonos nevű szerző-

tól. A költő és drámaíró Shakespeare-en kívül volt még egy vagy két William Shakespeare nevű szerző, nem valószínű azonban, hogy ezek sok könyvtári adatbázisban szerepelnének. Ezért a legtöbb rekordban nem fogják feltüntetni a születési évet, ha azonban az adatbázis több forrásból használ fel rekordokat, akkor bizonyosan lesznek olyanok is, ahol ezt feltüntetik, ezek a rekordok pedig új sorozatot alkotnak. Így ahelyett, hogy egy helyen megtalálhatnánk Shakespeare összes műveit, például a címek betűrendjében, három vagy négy külön sorozatot fogunk találni. Nyilvánvaló, hogy ez az a második eset, ahol a katalóguskészítőnek át kell szerkesztenie a rekordot, mielőtt bevinné az adatbázisba. A Reference Manual-ben van egy *Search Name* (keresőnév) elnevezésű adatelem, éppen ennek a problémának a megoldására. Ez az adatelem tartalmazza a kulcs-nevet, tehát a névnek azt a részét, amely szerint a listákban szerepelhet, valamint két kezdőbetűt. Ezt az adatelemet a Four Ways Group [11] kérésére vették fel a Reference Manual-be. A referáló és indexelő szolgáltatásokban sokkal inkább szükséges ez az adatelem, mint a könyvtári rekordokban, mivel az előbbieken a nevek közlés-módja kevésbé következetes.

Azt gondolhatnánk, hogy ez a probléma megoldható lenne egy általánosan használt, közös katalógizálási szabályzat létrehozásával. Nem lehetünk azonban ennyire optimisták! Mind a British Library, mind a Library of Congress az *Anglo-American Cataloguing Rules 2.* kiadását használja. Mivel azonban egyes dolgok különbözőképpen értelmezhetőek, más-más gyakorlat alakult ki a két intézménynél olyan esetekben, ahol a szabályok bizonyos fokú egyéni döntést tesznek lehetővé. Eltérések származnak abból is, hogy a két könyvtár gyűjteménye különböző, mindkét könyvtár katalógusaiban más és más szerzőket kell születési évük alapján megkülönböztetni, továbbá a nevek közlés-módja nincsen szabványosítva a két intézmény között. Ez a gyakorlatban azt jelenti, hogy ha a Library of Congress rekordjait a British Library katalógusaiba kívánják beépíteni, akkor semmi biztosíték nincs arra, hogy az egyéni és testületi szerzők nevének közlés-módja megegyezik, ezért az egységes katalógusszerkezet kialakítása érdekében minden egyes rekordot manuálisan szerkeszteni kell.

2.3.4 A többpéldányok kiszűrése

A rekordok cseréje során igen nagy problémát jelentenek a több példányban előforduló azonos rekordok. Az egy intézmény kezelésében lévő, de több forrásból származó rekordokat tartalmazó nagy adatbázisok esetében a többpéldányok viszonylag egyszerű algoritmus segítségével kiszűrhetőek. A British Library BLAISE MARC adatbázisa a British Library-ben készült rekordokon kívül a Library of Congress rekordjait is tartalmazza, melyek közül néhány a Kanadai Nemzeti Könyvtárban (National Library of Canada) készült. A Library of

Congress adatbázisából átvett rekordok közül kivesszük a brit kiadványok rekordjait, mivel ezeknek szerepelniük kell a helyben készült rekordok között. Vannak azonban olyan kiadványok, melyek kiadójának mindkét országban van kiadóhivatala. Ezért a következő szűrés az ISBN alapján történik, oly módon, hogy az új rekordok ISBN-jét összehasonlítják a fájlban meglévő minden egyes előzetes ISBN-nel. Így még mindig maradhatnak többes rekordok a fájlban olyan esetekben, amikor a kiadó ugyanannak a műnek más-más ISBN-t ad attól függően, hogy Angliában vagy az USA-ban terjeszti. A többes példányok azonban nem okoznak súlyos gondot a BLAISE MARC adatbázisban. Először is ugyanaz a rekord nem fordulhat elő kettőnél többször, hiszen túlnyomó többségük mindössze két helyről, a Library of Congress-ből vagy a British Library-ből származik. Másodszer, ha akár visszakeresés, akár katalógusba való beépítés céljából hívják le a rekordokat, egyik esetben sem okoz túl nagy gondot, ha egy helyett két azonos rekordot kapnak. A problémák akkor válnak súlyosakká, ha sok intézmény ad be rekordokat egyetlen adatbázisba. Egy igazi, osztott feldolgozáson alapuló együttműködésben, ahol a rekordok több forrásból származnak és a visszakeresés célja a leőhely azonosítása, egy rekordból csak egynek szabad lennie. Ellenkező esetben előfordulhat, hogy a kereső a rekord mellett csak egy leőhelyet talál, mert a többi leőhely egy másik, ehhez hasonló vagy ezzel azonos rekordnál szerepel.

M. E. Williams és K. D. MacLauray leírták egy, a többes példányok kiszűrésére szolgáló rendszer prototípusát [12], amelyet többek között az OCLC adatbázisban is kipróbáltak. Egy másik terület, ahol a rekordok összehasonlítása fontos, a *folyóiratcikkek* rekordjainak cseréje. A könyvek és folyóiratok esetében, különösen az újabbaknál, az ISBN vagy ISSN összehasonlításával a problémák többsége megoldható, a folyóiratcikkeknek azonban jelenleg még nincsen egyedi azonosító jelzete, (bár az ISO-ban már folynak munkálatok a bibliográfiai azonosítás kidolgozására [13]). Nagy számú többes példányhoz vezethet továbbá az is, ha a rekordok készítésében több intézmény vesz részt. Minden egyes intézmény általában egy-egy tárgykörért (tématerületért) felelős, az egyes tématerületek között azonban jelentős átfedés lehet, azon túl, hogy sok folyóiratcikk egyszerre több diszciplína szempontjából is relevánsnak bizonyulhat, pl. egy bizonyos rákellenes hatóanyagról szóló cikk könnyen előfordulhat kémiai, gyógyszerészeti, orvostudományi és egyéb, e tudományágakhoz kapcsolódó tárgyú adatbázisban is. Mivel növekvő azoknak a rendszereknek a száma, amelyek céljaira az UNISIST Reference Manual-t létrehozták, — beleértve a referáló és indexelő szolgáltatások közötti nemzetközi adatcserét is — a rekordok összehasonlítása egyre fontosabbá válik, éppen a források sokfélesége miatt. Az adatbázisba adott minden egyes rekord fontosabb adatmezőinek összeveté-

se az adatbázis összes rekordjával azonban célszerűtlen és gazdaságtalan lehet, mivel a számítógép viszonylag hosszú idő alatt tudja e műveletet elvégezni. Emellett sok többes példány így is átesne a rostán, mivel a különféle katalogizálási szabályok és az egyes intézmények eltérő gyakorlata következtében a rekordok a számítógép számára könnyen különbözőeknek tűnnek, és csak az emberi szem képes felismerni, hogy valójában ugyanarra a dokumentumra vonatkoznak.

Az új technika sem sokat segíthet e gondok megoldásában. Az egyetlen remény az, hogy a jövőben a számítógépes feldolgozás olyan olcsóvá válik, hogy a rekordok mezőnkénti összehasonlításának hosszadalmas művelete megoldható lesz. Ezeket a programokat azonban meg is kell írni, az algoritmusokat tehát olyan, katalogizálásban és információkeresésben jártas szakembereknek kell elkészíteniük, akik jól ismerik a megoldandó problémát. Ezek után pedig még mindig lesznek olyan rekordok, amelyeket a számítógép különbözőnek lát, és amelyekről csak nagyon bonyolult algoritmus, vagy éppen az emberi agy segítségével állapítható csak meg, hogy voltaképpen többes példányok.

2.3.5 A kiinduló formátum sajátosságaiból eredő különbségek

Egy másik ok, amelynek eredményeképpen az azonos dokumentumra vonatkozó rekordok különbözőnek tűnhetnek, még akkor is, ha az alkalmazott katalogizálási szabályok nagyon hasonlóak, éppen a csereformátumra való átalakítás. Az UNIMARC vizsgálata során [14] számos olyan eset került napvilágra, mikor egyik vagy másik kiinduló formátum egy-egy területen nem volt eléggé specifikus ahhoz, hogy megadja az adatokhoz tartozó kódokat, mint például annak a feltüntetését, hogy egy bizonyos szerző neve a bibliográfiai tétel besorolási eleme-e, társszerző-e, vagy éppen másodlagos beorolási elem. A csereformátumok, így az UNIMARC is, gyakran használnak úgynevezett „file-character”-eket annak a jelzésére, hogy a csereformátumban szükséges kódok nincsenek meg vagy nem vehetők át a kiinduló formátumban közölt adatokból. A más forrásból vett, csereformátumban közölt rekord várható felhasználásától, illetve attól függően, hogy a fogadó intézmény mennyire igényes a rekordok tökéletességére, igen sok munkába kerülhet a cserélt rekordok manuális korrigálása.

2.3.6 A fájlok belső egységességének biztosítása

Hosszan szóltam a többes rekordok kiszűréséről, és utaltam az új technikára. A címből már látható, hogy miről is lesz a továbbiakban szó. Eredetileg a fejezet címe az új technika és a csereformátumok kapcsolata volt, úgy döntöttem azonban, hogy az új technikáról csak a szerkesztésre alkalmas terminálok vonatkozásában lesz szó. A duplumok kiszűrése kifejezetten statisztikai

feladat, amelyről nálam sokkal hozzáértőbbeknek kell szólniuk. Akárhogy van is, nem hiszem, hogy találtak volna megoldást. A másik probléma, amiről szó lesz, nevezetesen a fájlok belső egységessége, azonban olyan természetű, hogy általában a csereformátumok használata során merül fel és kapcsolatos a formátum kidolgozásának módjával. Mivel a csereformátumok gyakran már meglévő formátumok közötti kompromisszum eredményeként jönnek létre, azért az adatelemek, definíciók, illetve a forma vonatkozásában nem mindig egybeváogók. Ezenkívül, mint már említettük, a különböző katalógizálási szabályok eredménye az eltérő értelmezésű feldolgozás. Mit lehet hát tenni? Kétféle lehetőség közül választhatunk:

- nem törődünk azzal, hogy szabványosak-e az adatok; ha az összevont fájlból nyomtatott jegyzéket készítünk, az minden bizonnyal rendezetlen lesz, lesznek benne ismétlődő nevek és hozzáférési pontok, továbbá következtlenül elkészített rekordok (ha ezt a fájlt online keresésre használjuk, akkor nem lesz sok problémánk, mivel az online keresésnél más megszkott dolog a gyengébb minőségű output és a képernyőn nem is látható egyszerre sok rekord, emellett a keresési stratégia kompenzálhatja az adatok hiányosságait);
- az adatokat szerkesztjük, mielőtt bevinnénk az adatbázisba. Ez a művelet költséges, itt azonban segítségünkre lehet a technika, és éppen ez a cikk második részének a tárgya.

3. Intelligens terminálok

Az intelligens terminálok olyan eszközök, amelyek alkalmasak távoli adatbázisok adatainak fogadására és egyidejű feldolgozására (bővebben ld. *Noerr*-nál [15]). A terminál intelligenciája bizonyos mennyiségű számítógép-memóriának köszönhető, amelyekben az adatok ilyen vagy olyan formában való manipulálására szolgáló programok tárolhatók. Ha a rekordokat adatbázisból veszik át, akkor rendszerint „új sor” karaktereket használnak az olvashatóság megkönnyítése érdekében. Közölni tudják továbbá a számítógéppel a terminál képernyőjének méretét, így a számítógéppel való kapcsolatteremtés és a password (jelszó) közlése után a gép tudni fogja, hogy milyen formában kell továbbítani a rekordokat úgy, hogy azok a terminál képernyőjén elférjenek. Az intelligens terminál úgy is programozható, hogy a rekordon további formalizálást végezzen. Ezen felül képes arra is, hogy kiszűrjön bizonyos, a felhasználó számára érdektelen adatmezőket, így ezek nem foglalnak helyet a képernyőn. Olvashatóbbá tudja tenni a rekordokat azáltal, hogy az egyes adatmezőkben a felhasználó kívánsága szerint megváltoztatja az online üzemmódban automatikusan közölt központosítási jele-

ket. A központi számítógép, amellyel a terminál kapcsolatban áll, szintén képes ezekre a műveletekre, ez a megoldás azonban drágább és sokkal kevésbé rugalmas. Az intelligens terminálok alkalmasak továbbá arra is, hogy „felhívják” az adatbázist tároló számítógépet, telefonvonalon közölgék a password-öt anélkül, hogy a kezelőnek a terminál bekapcsolásán és a kapcsolatteremtés létrehozására szolgáló gomb benyomásán kívül bármit is tennie kellene.

3.1 Mi is valójában az intelligens terminál?

Az intelligens terminál alkalmazása ma még nem széles körű. Fő alkotóelemük az integrált áramkör, melyet mikroprocesszornak vagy chip-nek is neveznek. Ebben a hardver-eszközbe számítógépi programokat, illetve karakterkódokat lehet betáplálni annak jelzésére, hogy a képernyőt irányító memóriarészéből kapott információ hatására milyen karakternek kell a képernyőn megjelennie. A chip-ek használata egyre elterjedtebbé válik. Az intelligens terminálokban levő mikroprocesszor hasonló a mikroszámítógépek mikroprocesszoraihoz. Mi tehát a különbség a mikroszámítógép és az intelligens terminál között? Az egyetlen különbség az, hogy az intelligens terminál egy terminál, tehát képes arra, hogy adatokat fogadjon távoli számítógépekből. Emellett vagy tudja mindazokat a műveleteket, amelyekre a mikroszámítógép képes, vagy nem, manapság azonban egyre inkább tudja. Valójában a legtöbb mikroszámítógéphez különféle eszközök (pl. modem vagy akusztikus csatlakozó) illeszthetőek, amelyek segítségével a mikroszámítógép telefonvonalon kapcsolódhat egy távoli számítógéphez. A mikroprocesszor alapú intelligens terminálok egyre elterjedtebbé válnak, és ma már olyan intézmények számára is elérhetőek, amelyek nem tudnak hardverre nagyobb összegeket áldozni. *K. D. Wise*, *K. Chen* és *R. E. Yokely* [16] ismertetik, hogy milyen mértékben csökkent a mikroprocesszorok ára 1970-től kezdődően. Míg 1970-ben 200 bit információ került 1 dollárba, 1980-ban már 500, napjainkban pedig több mint 1000 bit ára 1 dollár. Hasonlóképpen csökkent a mikroprocesszor alapú intelligens terminálok alkotórészeinek ára is, ugyanakkor, amikor az egyéb áruké mintegy 10%-kal emelkedett.

3.2 Az intelligens terminálok alkalmazása rekordok cseréjében

Mint már korábban említettük, a különböző adatbázisokból vett bibliográfiai rekordok egységesítésének problémája két okból következik, egyrészt az adatelemazonosítók, másrészt az adatok különbözőségeiből. Az intelligens terminálok mindkét probléma megoldására használhatók.

3.2.1 Példák

Példaként a British Library-ben kifejlesztett, a BLAISE adatbázisból online rekordok átvételére szolgáló rendszert szeretném bemutatni. Sajnálatos módon ez a rendszer sokakat csak elméletileg érdekelhet, hiszen kizárólagosan az Egyesült Királyság, illetve az Ír Köztársaság területén belül hozzáférhető. A BLAISE-CORTEX nevű rendszer Zentec ZMS-70 mikroszámítógépen üzemel, amelynek 64 Kbyte memóriája van, továbbá két lemez meghajtó négyeszeres sűrűségű 5 1/4 inch-es hajlékony mágneslemezhez (floppy disk), amellyel kb. 500 Kbyte tárolható. A rendszerhez 15 inch szélességű, 25 soros (egyenként 80 karakter) képernyő tartozik.

A rendszert brit és ír könyvtárak használják a BLAISE adatbázis rekordjainak átvételére és azok szerkesztésére. A könyvtár a rekordokat át tudja vinni hajlékony mágneslemezre, és vagy ezt küldi el a British Library-hez a célból, hogy a rekordokat offline üzemmódban beilleszék saját fájljukba úgy, hogy ezek a rekordok a következő katalógusban megjelenjenek, vagy pedig a lemez tartalmát telefonvonalon közvetlenül továbbítják a British Library fájljába. Az intelligens terminálok általam javasolt felhasználása a hívójelek konvertálása egyik formátumról a másikra. Az a bizonyos rendszer, amelyről szó van, általában nem használta az adatelemazonosítók konvertálására, mivel a British Library nagy számítógépen átalakítja a Library of Congress US MARC formátumban adott rekordjait UK MARC formátumra, mielőtt azok bekerülnének a BLAISE adatbázisba. A rendszer azonban használható erre a célra is, és minden bizonnyal vannak is olyan esetek, amikor a felhasználók meg kívánják változtatni a hívójeleket saját adatbázisaik számára.

3.2.2 Munkaműveletek

Az alábbiakban röviden ismertetjük, hogyan használható a rendszer idegen rekordok saját adatbázisába való bevételéhez. Tegyük fel, hogy katalogizálók vagyunk egy könyvtárban, és előttünk van egy nagy halom könyv, amelyet be kell vinnünk az adatbázisba. Mindenekelőtt ki kell keresnünk ezeket a rekordokat, azért visszük a rendszerbe minden egyes könyv ISBN-jét, illetve szabványos nemzetközi azonosító számítás hiányában a szerző nevét és a mű címét, mégpedig az adatbázist tároló számítógép által használt utasításnyelven. A telefonvonal és a számítógép gazdaságosabb használata érdekében célszerű a keresést előzetesen elkészíteni és a terminálon tárolni, majd egy gomb benyomásával a programot a számítógéphez továbbítani.

Valószínű, hogy nem minden rekordot fogunk megtalálni, de reméljük, hogy néhányat bizonyosan. Ezek a rekordok sorrendben megjelennek a képernyőn, majd a gép lemezre másolja őket. Ekkor egyenként átnézhetjük a rekordokat, elvégezhetjük a szükséges korrekciókat,

amelyekhez mindig hozzátartozik a könyvtár raktári jelzeteinek a feltüntetése. A képernyőn megjelent rekord külalakját az úgynevezett BLAISE-editor határozza meg. Minden adatmező egy három számjegyű hívójellel kezdődik, ezt követi két indikátor, egy szintjelző (szintkód) és egy ismételt hívójel. Egy ún. szűrő is beállítható, melynek használatával csak bizonyos mezők jelennek meg a képernyőn. Ha például adatbázisunkban soha nincs szükség a Dewey Tizedes Osztályozás jelzetére, ki kell szűrünk a 081 számú mezőt.

Az általam ismertetett rendszerhez hasonló rendszerek általában csak rövid felhasználói programok és szubrutinok alkalmazását teszik lehetővé, bizonyos speciális követelmények kielégítésére. A rendszerben általában speciális funkcionális billentyűket használnak, és amint a rekordot visszakeresték, a kezelő a kurzorral megkeresi a rekord első karakterét, pl. a kötelező 001 mező hívójelét, benyomja a megfelelő funkciók billentyűt, és egy mozdulattal több változtatást tud elvégezni az előre programozott rekordban. A hívójelek más jelekkel is helyettesíthetők. Az általam ismertetett rendszer használói ezt általában nem teszik, mivel ők is UK MARC formátumot használnak és a BLAISE adatbázisból is UK MARC formátumban közölt rekordokat kapnak. Az egyes adatmezőkben az almezők kódjai megfelelő központosítási jelekkel helyettesíthetők. A központosítási jelek bárhová beilleszthetők. A 300-as mezőben pl. az „oldal” jelentésű „p” után pontnak kell lennie, ezt azonban a British Library elhagyja. A pont automatikusan pótolható. Amint az automatikus változtatásokat a rekordban elvégezték, új rekord jelenik meg a képernyőn, a kurzor pedig az első sor elejére áll.

A rendszer lehetőséget nyújt a fájlban belüli bármely szövegrész keresésére is. Ha például a „Bibl.”; rövidítés helyett azt akarjuk írni, hogy „Bibliográfia”, meg kell keresnünk a „Bibl.” szövegrészt, és így minden mező, ahol ez a rövidítés szerepel, meg fog jelenni a képernyőn. Ezután elvégezhetjük a javítást, és az adatmező a javított formában fog szerepelni a fájlban.

Egy bizonyos billentyű lenyomásával új adatokat is vihetünk a rekordokba. Ha olyan mezőre van szükségünk, amely a kiinduló formátumban nem szerepel, például raktári jelzet, a program jelzi, hogy adatbevitelre vár, ellenkező esetben nem engedi, hogy a kurzor a kérdéses mezőről elmozduljon.

Ha a kezelő meggyőződött arról, hogy egy mező helyes, a megfelelő funkcionális billentyű lenyomásával a következő mező alá állíthatja a kurzort, amely azt jelzi, hogy a kérdéses mezőt a gép elfogadta. Más megoldásként az egyes mezők a képernyőn is javíthatók, ha pl. a beosorolási adatként szereplő nevet kell megváltoztatni, hogy azonos legyen a fájlban szereplő névalakkal, a kezelő az új szöveget a régi fölé írja, majd a megfelelő billentyű lenyomásával beviszi az adatmezőbe.

3.3 Az intelligens terminálok használatának összefoglalása

Összefoglalva elmondhatjuk, hogy a rekordokban szükséges mindenféle változtatás elvégezhető intelligens terminálok segítségével. A rendszeresen ismétlődő változtatások program segítségével automatikusan elvégezhetőek, egyéb változtatásokat pedig a rekordok áttekintése során lehet végrehajtani. A javított rekordok vagy egy hajlékony mágneslemezen tárolt fájlba, vagy az adatbázisba kerülhetnek, mely szolgálhat nyomtatott katalógusok készítésére, valamint online keresésre is. A cserereformátumban kapott rekordokban leggyakrabban szükséges változtatások típusait az alábbiakban foglaljuk össze:

- a) *Az adatelemazonosítók konverziója* (beleértve a hívójeleket, indikátorokat és almezőazonosítókat). Ez a művelet a legtöbb esetben automatikusan végezhető, bár az egymásnak nem pontosan megfelelő hívójelek esetében a gépnek jeleznie kell, hogy manuális beavatkozásra van szükség.
- b) *A besorolási elemek behelyettesítése*. Az adatelemek kompatibilitásától és a rekordok forrásától függően szükséges lehet az összes besorolási adat ellenőrzése.
- c) *A kódolt adatok kiegészítése*. Az egyes formátumok gyakran eltérő módon kezelik a kódolt adatokat. Ezek többsége automatikusan konvertálható, vannak azonban olyanok, melyeknél manuális beavatkozásra van szükség. Ha a kiinduló formátumban nem jelölik, hogy melyik a főtételek besorolási adata, a fogadó formátumban pedig szükséges ennek a jelölése, akkor ezt a legtöbb esetben manuálisan kell elvégezni.
- d) *Központozási jelek és besorolásra vonatkozó jelek*. Ezek a jelek esetenként automatikusan, esetenként a kezelő által, vagy a kettő kombinációjával vihetők be.
- e) *Hiányzó adatelemek pótlása*. Ezek közül egyesek nagyon fontosak lehetnek a rekord felhasználói számára, az intelligens terminálok szerkesztői programjával könnyen pótolhatók.

4. Következtetések

E cikk megírására azok a viták készítek, amelyeket éveken át folytattunk annak eldöntésére, hogy vajon valóban gazdaságos-e más intézmények által készített rekordok átvétele és beépítése a saját adatbázisba, vagy csak abban az esetben, ha a rekordok minimális változtatással pl. raktári jelzet feltüntetésével „honosíthatók”.

Megpróbáltam érzékeltetni, hogyan használhatók fel az intelligens terminálok e műveletek meggyorsítására, és hogyan teszik lehetővé, hogy a kezelő láthassa is azt a rekordot, amelyet bevisz az adatbázisba, ahelyett, hogy a javításokat távolból kellene végeznie. Inkább azzal foglalkoztam, ami lehetséges, nem pedig azzal, ami ma feltétlenül használatos. Sajnálatos módon eddig még

nincsenek számadataink arról, hogy mennyi munkaidő szükséges a rekordok itt ismertetett korrigálásához, így nem tudjuk összehasonlítani a manuális katalóguskészítés munkaidőigényével. Nagyjából nyilvánvaló azonban, hogy az adatok nagyfokú egységességét megkövetelő adattárakban mindig olcsóbb, ha a rekordokat helyben készítik. Mivel azonban a hardver-költségek csökkennek, és lesznek olyan intézmények, amelyek rendelkeznek az e feladatok ellátására kidolgozott szoftver használatára alkalmas hardverrel, vagyok annyira optimista, hogy azt gondoljam, az intelligens termináloknak a jövőben jelentős szerepe lesz a rekordok nemzetközi cseréjében a különböző országok különböző adatbázisai között, ezzel is elősegítve az információ szabad áramlását az egyes országok között, kiszélesítve az információcsere lehetőségeit és biztosítva a fejlődéshez szükséges információk tárolását és felhasználását. Ezek az Unesco Általános Információs Programjának az 1984–1989. évi középtávú terv tervezetében meghatározott főbb céljai [17]. Végül szeretném kifejezni azt a reményemet, hogy az intelligens terminálok egyre szélesebb körben fognak a világon rendelkezésre állni, éppen e célok elérésének megkönnyítésére.

Hivatkozások és jegyzetek

1. CLARKE, T. C.: Knowing your universals: UAP in relation to UBC = IFLA Journal, 4. köt. 1978. p. 129–133.
2. AVRAM, H. D.: MARC: its history and implications. Washington, Library of Congress, 1975.
3. AVRAM, H. D.: Various aspects of exchange formats: ISO 2709, UNIMARC, the Reference Manual, ISBD. In: Dierickx, H.: Towards a common bibliographic exchange format. Proceedings. . . Bp. OMKDK, 1978. p. 135–146.
4. UNIMARC: Universal MARC format. 2. átdolg. kiad. London, IFLA Office for UBC, 1980. p. V.
5. ISO 2709–1973. Format for bibliographic information interchange on magnetic tape. Geneva, ISO, 1973. Átdolg. kiad. 1981.
6. UNIMARC. Hivatkozott mű
7. DIERICKX, H.: Reference manual for machine readable bibliographic descriptions. 2. átdolg. kiad. Paris, Unesco, 1981. 346 p.
8. SIMMONS, P.: A common communication format for bibliographic exchange = Unesco Journal of Information Science, Librarianship and Archives Administration, 3. köt. 1981. 3. sz. p. 161–164.
9. Recommendations of the Symposium. In: Dierickx, H.: Towards a common bibliographic exchange format. Proceedings. . . Bp. OMKDK, 1978. p. 209.
10. UNIMARC. Hivatkozott mű. p. IX.
11. A *Four Ways Group* az ASIDIC, az EUSIDIC, az ICSU-AB és az NFAIS képviselőiből állt, és ezt a kérést egy kiadatlan dokumentáció tartalmazta, amelyet egy 1977 júniusában lezajlott értekezletre készítettek.

12. WILLIAMS, M. E.–MacLAURY, K. D.: Automatic merging of monographic data bases. Identification of duplicate records in multiple files: the IUCS scheme = *Journal of Library Automation*, 12. köt. 2. sz. p. 156–168.
13. ISO Recommendation: Bibliographic strip ISO. Geneva, 1956. (R/30). Átdolg. alatt.
14. A Deutsche Bibliothek vizsgálatot végzett, amelynek során különböző nemzeti könyvtárak UNIMARC formátumra konvertált rekordjait hasonlították össze. Az eredményeket hivatalosan még nem publikálták.
15. NOERR, P.: Intelligent terminals: some possibilities. In: *Hierarchical relationship in bibliographic descriptions*. Essen, Gesamthochschulbibliothek, 1981. (INTERMARC Seminar, 4.)
16. WISE, K. D.–CHEN, K.–YOKELY, R. E.: *Microcomputers: a technology forecast and assessment to the year 2000*. New York, Wiley, 1980.
17. Előzetes középtávú terv (1984–1989). 2. rész: *Information systems and access to knowledge*. Unesco General Conference, 4. rendkívüli ülés. Paris, Unesco, 1982. 7014ff bek.

Fordította: Szabó Andrea

HOPKINSON, A.: Nemzetközileg hozzáférhető adatbázisok rekordjainak szerkesztése intelligens terminálokon

Az elmúlt 15 évben jelentősen megnövekedett a bibliográfiai rekordok intézmények közötti cseréje. Ez nagy mértékben elősegítette az információs technika (information technology) fejlődése. Az adatbázisok közötti adatsere céljaira a könyvtárak az UNIMARC-ot, a referáló és indexelő szolgáltatások pedig az UNISIST Reference Manual-t használják. Az Unesco támogatásával kidolgozták a Közös Adatsere Formátumot (Common Communication Format, CCF) azzal a céllal, hogy áthidalja a könyvtárak és a referáló és indexelő szolgáltatások közötti különbségeket. A különböző intézmények által szolgáltatott rekordok közötti különbségek azonban oda vezethetnek, hogy a különféle forrásokból származó rekordokat nem lehet, vagy csak nehézségek árán lehet egy egységes adatbázisban felhasználni, még akkor is, ha azonos csereformátumra konvertálják azokat. A különféle forrásokból származó rekordok közötti nagyobb kompatibilitás biztosításának egyik módja, hogy a rekordokat rögtön beérkezésük után egységes formára szerkesztik. Ez a művelet időigényes és gazdaságtalanná is teheti a külső forrásokból vett rekordok felhasználását. Az intelligens terminálok alkalmazása e célra már praktikusabb megoldás lehet, mivel ezek képesek az adatbázissal való közvetlen kapcsolatra. A kívánt rekord lehívása után a megfelelő változtatások elvégezhetők akár az operátor közreműködésével, akár olyan program segítségével, amely megkeresi az egyes adatmezőket, és ott megváltoztatja a hívójelet vagy bizonyos adatokat, mielőtt a rekordot az adatbázis fájljához rendelnék. Az intelligens terminálok alapját képező mikroprocesszorok árának csökkenésével ez a művelet a jövőben egyre elterjedtebbé válik.

HOPKINSON, A.: Editing records from internationally available databases on intelligent terminals

The last 15 years have seen a tremendous growth in the exchange of bibliographic records between organizations, which has been assisted by advances in information technology. The library community has developed UNIMARC and the abstracting and indexing community are using the UNISIST Reference Manual to facilitate the transfer of bibliographic data between databases. Unesco is establishing a Common Communication Format in an attempt to bridge the gap between the library and the A & I community. However, different practices in record creation between organizations providing records mean that records from different sources cannot always be merged comfortably into one database, even if they have been converted into the same exchange format. One way of achieving greater compatibility between records from different databases is by editing the records as they are received. This can be time consuming and can make use of records from outside sources uneconomic. Intelligent terminals can make this more of a practical proposition. They can assist in dialling up the appropriate source database. When the desired record is obtained, changes can be made to it, either by the intervention of the operator or by programs in the terminal which can search for a particular field and make a change to the tag or to the data in a field before adding the record to the file, stored probably on a mainframe computer. This kind of operation is expected to increase in the future, as microprocessors which are the basis of intelligent terminals fall in price.

ХОПКИНСОН, А.: Редактирование на интеллигентных терминалах записей баз данных, доступных по международным линиям связи

За последние 15 лет обмен библиографическими записями между организациями значительно возрос. Этому содействовало развитие информационной технологии. В целях обмена библиографическими данными между базами данных, библиотечное общество внедрило коммуникационный формат UNIMARC, а реферирующая и индексирующая службы использовали коммуникационный формат, описанный в UNISIST Reference Manual. При поддержке ЮНЕСКО разработан Единый Коммуникационный Формат (Common Communication Format) с той целью, чтобы ликвидировать разницу между библиотеками и индексирующими и реферирующими службами. Опыт различных учреждений показал, что библиографические данные, полученные из разных источников не могут быть удобно объединены в одну базу данных даже в том случае, если они переведены в один и тот же коммуникационный формат. Возможным путем достижения лучшей совместимости записей, полученных из различных баз данных, является их редактирование в ходе приобретения. Это может потребовать много времени и, таким образом, сделать записи, полученные из внешних источников неэкономичными. Применение интеллигентных терминалов имеет большое практическое значение потому, что позволяет осуществлять связь непосредственно с базой данных. После вызова желаемой записи в нее могут быть внесены соответствующие изменения либо оператором, либо с помощью программы терминала, которая ведет поиск полей данных и изменяет там метку данного элемента или одно из определенных данных перед присоединением данной записи к файлу главной ЭВМ. При понижении цен на микропроцессоры использование интеллигентных терминалов будет находить все более широкое применение в будущем.

HOPKINSON, A.: Redigierung der Rekorde international zugänglicher Datenbasen mit intelligenten Terminale

Der Austausch bibliographischer Rekorde zwischen den Institutionen nahm, gefördert durch die Entwicklung der Informationstechnologie, in den letzten 15 Jahren beträchtlich zu. Für die Zwecke des Datenaustausches zwischen den Datenbasen benützen die Bibliotheken UNIMARC, die Referier- und Indexierdienste das UNISIST Reference Manual. Unterstützt durch UNESCO wurde ein gemeinsames Datenaustauschformat CCF (Common Communication Format) ausgearbeitet, um die Unterschiede zwischen den Bibliotheken und den Referier- und Indexierdiensten auszugleichen. Es können nämlich die zwischen den durch die unterschiedlichen Institutionen gebotenen Rekorden bestehenden Unterschiede dazu führen, dass die Rekorde aus unterschiedlichen Quellen nicht, oder nur mit grossen Schwierigkeiten benutzt werden können, auch wenn sie auf ein identisches Austauschformat konvertiert worden sind. Um eine grössere Kompatibilität zwischen den Rekorden aus unterschiedlichen Quellen zu sichern bietet sich als eine gute Methode an, die Rekorde sogleich nach ihrem Eintreffen in das einheitliche Format zu redigieren. Dieser Vorgang kann aber die Nutzung von Rekorden aus fremden Quellen zeitaufwendig und unwirtschaftlich gestalten. Für diesen Zweck erscheint der Einsatz intelligenter Terminale schon praktischer, da diese für einen direkten Kontakt mit der Datenbase geeignet sind. Nach dem Abruf des gewünschten Rekords können die nötigen Veränderungen durch das Mitwirken des Operatoren oder mit Hilfe eines solchen Programmes durchgeführt werden, welches die einzelnen Datenfelder aus sucht, dort das Rufsignal oder gewisse Daten verändert, bevor der Rekord dem File der Datenbase zugeordnet wird. Im Zuge der sinkenden Preise der die Grundlage der intelligenten Terminale bildenden Mikroprozessoren wird dieser Arbeitsprozess sich sicherlich in einem immer breiteren Kreis durchsetzen.