

Az osztrák irodalom virtuális könyvtára

Előzmények

Világszerte számos projekt foglalkozik a könyvek és folyóiratok visszamenőleges digitalizálásával. Ezeknek a vállalkozásoknak – propagandahatásuk mellett – fontos művelődés- és tudománypolitikai jelentőségük van, hiszen a jövőben a nyilvánosság csak azokat a dokumentumokat veszi tudomásul, amelyek elektronikus formában is rendelkezésre állnak. A világhálón még nagyon szerény a német irodalom, különösen pedig az osztrák irodalom aránya. Ezen a helyzeten kíván változtatni az ALO (*austrian literature online*) program, amely egyelőre a 18. sz. végétől 1930-ig megjelent, legfontosabb 1000 osztrák könyv digitalizálását tűzte ki célul.

1999-ben egy kivitelezhetőségi tanulmányt végeztek a könyvtárak és egyetemek munkaközössége kérésére a Tudományos, Oktatási és Művelődési Minisztérium megbízásából: megvizsgálták egy osztrák virtuális könyvtár megvalósíthatóságát, és ajánlásokat dolgoztak ki. Ennek keretében felállították egy ilyen könyvtár prototípusát, amely képes a könyvek és folyóiratok digitális konvertálására, az elektronikus faksimilék, az automatikusan rögzített teljes szövegek és a metaadatok tárolására, a faksimilékben, a teljes szövegekben és az indexekben keresési műveletek végrehajtására, különféle elektronikus output formák (beleértve a vakok számára audio-adatsorokat) előállítására, egy-egy könyv integrált reprintelésére, az ALEPH hálózati központi katalógusban a digitális könyvek szerepeltetésére.

A prototípus után sor kerülhet a virtuális könyvtár felállítására. Ehhez célszerű lesz minél több szponzort megnyerni. Pl. az egyes tartományokat és városokat a helyi érdekű könyvek digitalizálásának támogatására, magánszemélyeket, intézményeket és vállalatokat pedig bizonyos témakörökbe eső anyagok finanszírozására.

A digitális konverzióval elő kell mozdítani a már működő szolgáltatások hatékonyságát is. Így pl. a könyvtárközi kölcsönzésben nehéz volt a régi anyagokhoz hozzáférni a könyvtárak vonakodása miatt; a digitalizált könyvbe most a weben keresztül lehet bepillantani. (A könyvtárközi kölcsönzésben felmerülő igények alapján lehetne kijelölni a digitalizálandó anyagot is.) A különféle kutatási programok keretében kellene megoldani a kon-

vertálás egy részét: a nagyobb mennyiségű kutatások esetében mindjárt a kezdet kezdetén digitalizálni kell a legszükségesebb anyagot, ami a későbbiek során nagy könnyebbséget jelentene.

Az ALO-rendszerrel szemben támasztott követelmények

Az ALO-rendszerben elektronikus könyveket állítanak elő, amelyeket egy digitális könyvtárban kezelnek, és az interneten keresztül bocsátják a használók rendelkezésére. Ezért maximálisan biztosítani kell az adatok tartósságát, s a tárolásnál csak szabványosított formátumokat szabad használni.

Az ALO-ban egy elektronikus könyv a következő elemekből áll: az eredeti könyvalak digitalizált faksimiléi, ezeknek az oldalaknak géppel olvasható szövege, valamint a könyv eredetijének és elektronikus verziójának metaadatai.

Az archiválást szolgáló faksimiléknek legalább 400 dpi felbontással és 24 bites színmélységgel kell készülniük. A HTML-böngészők számára azonban egy másikra is szükség van 100 dpi-vel és 8 bites mélységgel. A könyv szövegét vagy begépelik, vagy szövegfelismerő programmal rögzítik.

Az ALO-könyvtárhoz egy egyszerű menedzsment-rendszeren keresztül lehet hozzáférni. Ebben olyan programok vannak, amelyek elősegítik a faksimilék és a teljes szövegek előállítását. A rendszeren keresztül lehet beilleszteni vagy kivonni könyveket. A használók a weben érik el a könyvtárat. A HTML-lapokon keresik meg a könyveket, amelyeknek faksimiléjét vagy szövegét különféle formátumban töltik le. Azok a használók, akik előnyben részesítik a nyomtatott formát, valamely könyvszolgáltatótól (Book-On-Demand-Service) rendelhetik meg. A keresés a metaadatok (szerző, könyvcím, tárgyszavak) szerint vagy a teljes szövegben történhet.

Az ALO-rendszer megvalósítása

A szövegfelismerő (OCR) és a szkennelő rendszer külső programok, amelyek nincsenek az ALO-rendszerbe integrálva. Ezeknek a programoknak

az outputját a menedzsmentrendszer használja fel az elektronikus könyv előállítására. Az ALO-könyvtárban tárolt könyvhöz a használói interfészen keresztül lehet eljutni.

A gráci Egyetemi Könyvtárban egy Minolta-szkenner szolgálja ki az ALO-rendszert. A faksimilét GIF formátumba konvertálják, és az ALO-szerver merevlemezén tárolják. Tekintettel arra, hogy az 1942 előtt megjelent német könyvek 80%-a gót betűs, nagy gondot fordítottak a szövegfelismerő rendszer kiválasztására. A legjobbnak a *Finereader* bizonyult, 94,75%-os biztonsággal. (A *Recognita* csak 74,65%-ot produkált.)

Az ALO-könyvtár képezi a rendszer magját. Java nyelven implementálják, s ezért platformfüggetlenül lehet alkalmazni. A Java egyébként számos eszközzel rendelkezik XML-adatsorok feldolgozására, s servletek segítségével sok ismert web-szerverhez csatlakozik.

A könyvek szövegét egy *TEI*LITE formátumú adatbázisban tárolják. Ennek előnye, hogy emberek és számítógépek egyaránt olvashatják; elősegíti az adatbázis tartósságát, és megkönnyíti az adatok publikálását az interneten. A leíró metaadatok (szerző stb.) tárolhatók benne, az adminisztratív és strukturális metaadatok azonban, amelyek az elektronikus könyv felépítését és formátumát írják le, nem. Ezeket a MOAII adatformátumban tárolják. A faksimilét GIF formátumban tárolják, az ebben lévő adatok minden webböngészőn hozzáférhetők.

A menedzsmentrendszer az elektronikus könyvek előállítására és tárolására szolgál, s a következő részekből áll: *html2tei* (leolvassa a Finereader által előállított HTML adatsort, és egy TEI adatsorból konvertálja); *makemoa* (a TEI adatsorból és a faksimiléből automatikusan állít elő egy MOA adatsort a könyv számra); *addbook* (a TEI, MOS és faksimile adatsorokból álló elektronikus könyvet tárolja az ALO-könyvtárban); *delbook* (törli a könyvet az ALO-könyvtárból). Az ALO-könyvtár használói interfésze statikus és dinamikus HTML-oldalakkal áll.

A könyvtartalmak oldalait a Java-servletek dinamikusan állítják elő. A servletek teszik fel a kérdéseket az ALO-könyvtárnak, s az XML-választ XSL-Stylesheets segítségével HTML-oldalakra váltják.

A metaadatok (szerző, cím stb.) szerinti keresést ugyancsak egy servlet végzi, és listát készít a megtalált könyvekről. A teljes szövegben egy meglévő rendszerrel lehet keresni, amely a TEI-adatsorokból indexet készít, s kérdés esetén ebben az indexben keres. A keresés eredményeként egy listát ad azokról a könyvekről, amelyek a kérdéses szavakat tartalmazzák.

Az ALO projekt a könyvformátumú szolgáltatásra (Book-On-Demand-Service) a Libri céget választotta.

EGGER, Alexander-MÜHLBERGER, Günter: ALO oder die virtuelle Bibliothek der österreichischen Literatur. Ein Arbeitsbericht. = Bibliotheksdienst, 34. köt. 6. sz. 2000. p. 958-967./

(Papp István)

A gazdasági válság hatása a tudományos és műszaki információra Ázsia néhány országában

A fejlett és fejlődő, gazdag és szegény országok között a leglényegesebb különbség a tudás, az ismeretek terén levő szakadékban mutatkozik. Az ebből fakadó kihívás 1990 körül jelentkezett, amikor néhány országban a gazdaság két számjegyű fejlődési ütemet produkált, s másutt is erősnek és növekedőnek mutatkozott.

A gazdasági buborékot azonban nem lehetett tovább fűjni. A gazdasági világválság először Ázsiában jelentkezett 1996 júliusában, s hatása máig érződik. A kormányoknak felül kellett vizsgálniuk gazdaságpolitikájukat, az egyes intézmények pe-

dig szigorú megszorításokat szenvedtek el. Ez a legújabb kihívás a fejlődő országok számára, amelyek az információ terjesztésében eddig inkább a szóbeli hagyományozással éltek, s nem a nyomtatott papírral.

Az információs szolgálat egy ország gazdasági jólétének jelentős tényezője, s jelzi a fejlődő országokban is a gazdaság sikerességét vagy kudarcait. Néhány döntéshozó Ázsiában is felismerte ezt az összefüggést, s megkezdte a nemzeti információs politika erősítését. A válság azonban éppen akkor ütött be, amikor több ázsiai ország a