

Kép- és videóarchívumok – válasz a változó igényekre

Sopov István PhD hallgató
Budapesti Műszaki és Gazdaságtudományi Egyetem,
Műszaki Menedzsment Doktori Iskola

Kivonat

Kulturális adatok feldolgozásával foglalkozó szervezeteknek különböző és folyamatosan fejlődő platformokon kell megjeleníteniük anyagaikat, és ez az archívumok nagyfokú rugalmasságát követeli meg.

Ahogy újabb és újabb eszközök és platformok jelennek meg és ahogy változnak a felhasználói elvárások, úgy válik szükségessé az új és az archívumban szereplő régi anyagoknak új formátumokban történő megjelenítése. Az így vázolt igényekre adott sikeres válaszok egyikeként kerül bemutatásra a PORT.hu kép- és videóarchívuma, amely könnyen tud kiszolgálni a tervezéskor még ismeretlen formátumokra vonatkozó megjelenítési kéréseket is.

A megvalósított koncepció lényege, hogy az egyes objektumok több példányban kerülnek letárolásra, és ezek száma bármikor tovább bővíthető. Az előadás röviden bemutatja az adatbázis működését és az adatbázis tartalmának néhány megjelenési felületét a formátumokra vonatkozó érdekességek és az eddigi tapasztalatok kiemelésével.

Bevezetés

A számítástechnika fejlődésével megváltoztak a tudás megjelenési formái. A korábbi világképben egy tiszteletet ébresztő épületet képzeltünk el, ahol sok tízezer könyv található szigorú rendben. A hatalmas olvasótermekben kis asztali lámpák fénye mellett merülnek az új generációk a régiek által összegyűjtött és rendszerezett ismeretanyagba. Ez a kép változott meg az információs technológiák fejlődésével és térnyerésével. Az összegyűjtött tudást ma egyszerűen és viszonylag alacsony költségekkel lehet a digitális világon keresztül megosztani, melyhez az informatikai írástudók bármikor és bárhol könnyen hozzáférhetnek.

A megismerést és kutatást segítő könyvtárak először csupán katalógusaikat alakították át elektronikus katalógusokká, az utóbbi tíz évben azonban maguk a gyűjtemények is egyre több esetben érhetőek el digitális formában. Ez természetesen összhangban van a könyvtárak eredeti szellemiségével is, de a folyamatot más érdekek is segítik. Ezek közül én kiemelném, hogy a nemzetek versengésében az információs- vagy tudásalapú társadalom kiépítése lett a legfontosabb terület, melynek lényegi eleme és katalizátora lehet a kulturális örökség digitalizálása.

Minden ország jelentős erőforrásokat rendelt a nemzeti kultúrához tartozó értékek digitális környezetben történő megőrzéséhez és megosztásához.

A könyvtárak és más közintézmények az elektronikus információs környezetünknek azonban csak egy részét adják, melyben számos más szereplő is szolgáltatásokkal és fejlesztésekkel jelentkezik (cégek, non-profit szervezetek és magánemberek egyaránt). A különböző jogállású szereplők hasonló technikai részfeladatokkal találkozhatnak, melyeknek tanulságait érdekes lehet megosztani.

A kép- és videóarchívumok néhány általános jellemzője

A digitalizált gyűjtemények a szöveges dokumentumok mellett képi, hang- és mozgóképes anyagokat is egyre nagyobb számban tartalmaznak. Ezen a területen nehezebb hosszú távon előre látni, hogy mely szabványok lesznek életképesek és hosszú életűek, sőt még rengeteg technikai megoldásra lehet számítani a terület viszonylag fiatal volta miatt (pl. folyamatosan fejlődnek a katalogizálást automatizáló hang- és alakfelismerő alkalmazások).

Egy digitális kép- vagy videóarchívum létrehozásakor különböző kiindulási helyzetek állhatnak fenn. Egyik jellemző esetként az archívum létrejöhet egy digitalizálási projekt keretén belül, amikor egy már létező gyűjteményt ültetnek át digitális hordozóra (pl. papírképek, videoszalagok egy könyvtár birtokában). A másik véglet lehet, amikor egy választott nulla időponttól kezdve érkező új anyagokból épül fel a folyamatosan bővülő archívum (pl. köztéri reklámfigyelés, televíziós hírműsorok archiválása). Ez azért előnyös, mert ilyenkor általában lehetőség van a digitális archívumhoz leginkább alkalmas formátumban beszerezni az anyagokat (pl. digitális köteles példányok). Általános esetben az archívum felépülhet részben hagyományos és részben digitális formában levő anyagokból és bővítése is elképzelhető különböző formátumokban megküldött anyagok folyamatos feldolgozásával.

A képek digitalizálása történhet szkennelvel vagy digitális fényképezőgéppel. Az utóbbit általában akkor használják, ha az eredeti dokumentumok sérülésével járhatna a lapszkennelbe helyezésük. Ezek a megoldások természetesen a szöveges anyagok esetében is hasonlóan működnek. A videoanyagokat digitalizáló-kártyával rögzítik, ami az eddig említett hardver eszközökhöz hasonlóan a személyi számítógépek gyakori, így nem különösebben költséges tartozéka. Más a helyzet például celluloid szalagok digitalizálásakor, amelyhez filmszkennelre van szükség, ami az eddig felsorolt eszközöknél nagyobb beruházást jelent.

A digitális rögzítés után (capture), szükség lehet a dokumentumok szerkesztésére (edit), mielőtt azok végleges archívumbeli formájukat megkapnák. Ez jelentheti például a kontraszt és a színhelyesség beállítását vagy a nem kívánt részek eltávolítását. Módosításokra a

digitális formában megkapott vagy előállt anyagok esetében is szükség lehet.

Az anyagokból érdemes a mester példányon kívül egy kis méretű nézőképet (thumbnail) is rögzíteni, így a katalogizálás során és később a felhasználók képernyőin sem kell rögtön a nagyméretű képet megjeleníteni, csak ha a nézőkép megtekintése után is szükség van rá. Ezt videoanyagnál is érdemes megtenni egy vagy több pillanatképpel, de létrehozható egy gyorsan megnyitható alacsony felbontású verzió is. A mester példányoknál lehetőség szerint adatvesztés nélküli tömörítéses (pl. TIFF kép) vagy érzékelhető minőségromlás nélküli veszteséges tömörítéses formátumokat érdemes használni, míg a nézőképeknél előnyös, ha nagy tömegben sem igényelnek számottevő tárolókapacitást.

Az archiválandó dokumentum így előállt reprezentánsa kerül az archívumba, miután a megfelelő meta-adatok is rögzítésre kerültek. Az archívum által használt leíró meta-adatok szerkezetének kialakításánál mindenképpen érdemes a nemzetközi szabványokat és ajánlásokat követni az archívumok közötti könnyebb átjárhatóság és kommunikáció biztosítása érdekében.

Felhasználói igények

A fenntartók és felhasználók oldaláról az egyik alapvető igény a digitális megőrzés esetében a tárolás tartóssága és az információkhoz való hozzáférés (megjelenítés) korszerűségének fenntartása.

A digitális archívumokba létrehozóik jelentős energiákat fektetnek, mely után jogosan várják el, hogy azok hosszú ideig („örökké”) szolgálják a közönségüket. Itt érdemes idézni Jeff Rothenberg mondását, miszerint „a digitális információ örökké megmarad, vagy csak öt évig – akármelyik is következik be előbb”. Ez vonatkozik a digitális adathordozóknak a jó minőségű papírhoz viszonyítva rövid élettartamára és az adathordozók és fájlformátumok gyors evolúciójára, ez a 8-10 évvel ezelőtt eltárolt adatokhoz való egyszerű hozzáférést is költséges és időigényes „technorégészeti” projektté varázsolhat.

A felhasználók néhány év múlva már esetleg teljesen megváltozott eszközökön és más szoftverkörnyezetben kívánják az adatokat megtekinteni. Az új környezet a maitól jelentősen eltérő adatformátumot igényelhet, amelyet nehéz előre megjósolni, de felkészülni érdemes rá. Ez azt jelenti, hogy az adatbázisokat és a hozzájuk kapcsolódó szoftvereket úgy célszerű felépíteni, hogy külön programozás és átstrukturálás nélkül legyenek képesek a tárolt anyagok formátumainak bővítésére.

Így amikor az új formátumot a saját archívumunkban is bevezetjük, nem kell majd az adatszerkezetet és az alapvető eljárásokat megváltoztatni, csupán a konverziókat kell megoldanunk. Az új anyagok felvételekor ez a

digitalizálás és anyagbekérés módosulását jelentheti, vagy csupán a rögzítési művelet megváltoztatását. Az archívumban már szereplő tételeket pedig egy konverziós eljárással kell az új formátumra átültetni, amely például egy konvertáló program kötegelt meghívásával valósulhat meg.

A végfelhasználói igények másik fontos területe a kereshetőség. A digitalizált gyűjtemény csak akkor hasznosítható, ha a használói könnyen meg tudják találni benne az őket érdeklő tételeket. A katalogizálás egységesítésére a nyolcvanas évek második felében a MARC szabvány terjedt el a hagyományos dokumentumok esetében, míg az elektronikus formátumú dokumentumoknál a meta-adatok javasolt köre a Dublin Core adatkészletet veheti alapul. Az ajánlások és szabványok követésével a különböző gyűjtemények összekapcsolhatóvá válnak és egyszerre lehet bennük a kereséseket végrehajtani. Az ajánlások nem csupán a használt meta-adatok szerkezetére vonatkoznak, hanem a besorolásoknál használt osztályozásra is egységes szótárakat adnak meg.

Egy multimédiás adatbázis esetében a meta-adatok megadása és a kategorizálás időigényes feladat, de így válhat az archívum a végfelhasználók számára kiaknázható forrássá, míg ennek hiányában nehezen mobilizálható adattenger lesz csupán.

A kereshetőséghez szorosan kapcsolódik a kezelőfelület kérdése, mely különböző platformokról hasonló formában kell elérhető legyen. A kezelőfelületről lekérhetőek a digitális és nem-digitális dokumentumok adatai beleértve az összekapcsolt gyűjteményekre vonatkozó találatokat is.

A PORT.hu kép- és videóarchívumának bemutatása

A PORT.hu B2B és B2C tartalomszolgáltatással foglalkozik, elsődleges szakterülete a kulturális programinformáció. A kulturális információszolgáltatás szinte elképzelhetetlen a szöveges adatok mellett különböző képi anyagok nélkül. Ezek elsősorban személy-, jelenet- és helyszínfotók, esetleg teremvázlatok. A PORT adatbázisa a nyolc évvel ezelőtti megalkotásától kezdve tartalmazott képanyagokat, melyek az adatbázisban csak a megfelelő személyhez vagy műsorhoz (film, tv műsor, színházi előadás) kapcsolódtak technikai adataik (elérési út, magasság és szélesség adatok) és képaláírásuk megadásával.

Az internetes sávszélességek növekedésével a látogatók elvárásai is megváltoztak. Egyre több képet szeretnének megnézni, sőt a különböző formátumú videó állományok megtekintéséhez is csak pár másodpercet kell várakozniuk a szélessávú eléréssel rendelkező felhasználóknak, akiknek száma öröndetesen gyarapodik.


A cég üzleti partnereinek köre is bővülő szolgáltatási igényekkel jelentkezett (a tartalom újabb területeit kívánta készen beszerezni), illetve a technikai fejlődéssel új ügyfelek is megjelentek (például a mobilszolgáltatók is multimédiás tartalmat kezdtek biztosítani előfizetőiknek).

Ezek a tendenciák arra ösztönözték a céget, hogy különálló multimédia-adatbázist hozzon létre, és továbbfejlesztett módon kezelje a képi- video- és hanganyagokat. A megközelítés lényege, hogy a hozzáférhető legjobb minőségben beszerezett anyagokat azok minőségét megtartva egységes formátumban tárolja a rendszer, és tetszőleges számú másodpéldányt hoz létre a felhasználási igényeknek megfelelő paraméterekkel. A változtatás vonatkozhat például a formátumra, a felbontásra, a tömörítés módjára, a feliratozás módjára vagy a színmélységre. Az automatizálható átalakítások mellett természetesen lehetőség van szerkesztői munkát igénylő másodpéldányok létrehozására is (például a videoanyagból egy rövidebb anyag vágása, képből egy jellemző négyzet alakú részlet kiemelése, stb.). A másolatok részletgazdagsága természetesen legfeljebb csak az eredeti szintjét érheti el.

Az adatbázis a PORT.hu többi adatbázisához hasonlóan az Oracle adatbázis-kezelő 9.2.0.4 verzióját használja, a kapcsolódó szerkesztőségi rendszer különböző részei PL/SQL, Java és Visual Basic nyelveken készültek. Az anyagok tárolásának szempontjából két alapvető út kínálkozik: az adatbázisban tárolás és a fájlstruktúrában tárolás. Az adatbázisban tárolt kép- és videoanyagok esetén (az Oracle InterMedia modul által nyújtott lehetőség) jelentősen megnő az adatbázis mérete, de könnyen megőrizhető a konzisztencia (ne lehessen képet eldobni, amíg van rá hivatkozás). A fájlként tárolás esetében rendszeresen konzisztencia ellenőrző rutinokat kell futtatni a hivatkozások és a fájlok meglétének ellenőrzésére, de megfelelő archiválási rend és szerkesztőségi rendszer megléte esetén ugyanúgy fenntartható az archívum konzisztenciája. Azt tapasztaltuk, hogy a fájlként való tárolás esetén a megjelenítés kevésbé erőforrás-igényes. A PORT.hu adatbázisa esetén az egyik döntő szempont az volt, hogy a fájlként tárolt anyagokon könnyen le lehessen futtatni utólag is parancssorból indítható műveleteket (például tömeges átméretezés, vízjellel való ellátás, stb.).


A PORT.hu tevékenységéhez kapcsolódóan jellemzően eleve elektronikus formában beérkező képeket rögzít, melyek e-mailben vagy CD-n érkeznek. Az új felépítésű archívum feltöltésének első szakaszában a szerkesztők a korábban beérkezett képek eredeti példányait dolgozták fel. Ezt követően az archívumot hétről-hétre a televízió adók, mozik és színházak aktuális műsorához kapcsolódóan kellett csak bővíteni.

Képek [Teljes galéria \(10\)](#)




Klikkeljen a képekre a nagyításhoz!

Interaktív, 360 fokos panoráma



Klikkeljen a képekre a nagyításhoz!

Videók



Klikkeljen a képre a lejátszáshoz!

PORT.hu | [TV](#) | [Mozi](#) | [Koncert](#) | [Színház](#) | [Rádió](#) | [Étterem](#) | [artPORT](#) | [Játék](#)
[Impresszum](#) | [Médiaajánlat](#) | [Kapcsolat](#) | [Magunkról](#)

1. ábra A médiafájlokat bemutató részlet a Lánchíd adatlapjáról (PORT.hu)

A megrendelői igények és a költségvonzatok figyelembevételével az eredeti képeket legfeljebb 2000*2000 pixel méretben és kevés adatvesztéssel járó JPEG formátumban tárolják (érdeemes hozzátenni, hogy ritkán érkezik be ennél nagyobb kép). Az eredeti fájlokat csak a további formátumok legyártására és a nyomtatott média kiszolgálására használják.

A képek feldolgozására szolgáló szerkesztőségi rendszer az eredeti kép rögzítésekor automatikusan legyártja a további szükséges formátumokat. Jelenleg ezek további JPEG (erősebb tömörítéssel készített 400, 300, 250, 180 és 75 pixel méretű változatok), WBMP (wireless bitmap) és GIF formátumok (fix pixel méretű és legfeljebb 2kB fájl méretű változat). A JPEG képek különböző internetes megjelenésekre alkalmasak és a digitális tévéadásokat kísérő EPG (Electronic Programme Guide) és super teletext alkalmazásokban jeleníthetőek majd meg. A WBMP és GIF formátumokat mobiltelefonokra szánt tartalomként kerülnek felhasználásra, ahol a sávszélességgel való spórolás elsődleges szempont. A PORT.hu képadatbázisa többek között az [origo], a kultúra.hu és a Pannon GSM wap portáljának oldalait szolgálja ki.

Az előzőekkel azonos rendszerben történik a videoanyagok feldolgozása is, ahol nagyobb szerep jut a digitalizálásnak. A saját készítésű anyagokat a felvételkor is használt DV formátumban archiváljuk, míg a videoszalagon vagy elektronikus formában beérkező anyagokról tömörítetlen AVI

formátumban (HuffYUV kodek) készül el az elsődleges digitális példány. Ez nagyméretű, de gyorsan szerkeszthető és konvertálható fájlokat eredményez. Az állományok méretéből adódóan az eredeti példányokat csak DVD-re írva tároljuk, míg a másodpéldányok azonnal hozzáférhetőek a szerkesztőségi rendszerből. A másodpéldányok jelenleg hat formátumban kerülnek legyártásra a megrendelői igényeknek megfelelően. Az internetre készített állományok egyszerre tartalmaznak több sávszélességre optimalizált változatokat.

	Webes felületekre	Mobiltelefonra
RealVideo	240*180 pixel	176*144 pixel
Windows Media Video	320*240 pixel	176*144 pixel
3GP Video	-	128*96 pixel
		176*144 pixel

1. Táblázat Kimeneti videoformátumok a PORT.hu adatbázisában

A felhasználói igények változását jól lehet követni a formátumok használatában. Kezdetben csak RealVideo formátumban voltak elérhetőek az anyagok, amelynek lejátszásához külön szoftver letöltésére volt szükség, és a streamelés segítségével kisebb sávszélességen is élvezhető minőségben jelent meg a videó. Ehhez az időszakhoz kapcsolódik a PORT azon fejlesztése, mely a kép alatt külön mezőben jelenítette meg a feliratokat szövegfájlból beolvasva, így azok jól olvashatóak maradtak. A szélessávú internet elterjedésével és a Windows operációs rendszerekbe beépített Windows Media Player fejlődésével annak saját WMV formátuma került előtérbe. A felhasználóknak már nem kellett sokat várniuk a teljes anyag letöltéséhez, ezért csökkent a streamelés jelentősége (bár azt saját formátumban a Windows Media Player is támogatja) és a sávszélesség növekedéséből adódóan tisztább és nagyobb méretű képeken a képbe konvertált feliratok sem mosódnak el. A GPRS megjelenésével a mobil eszközökre is felkerültek a videólejátszó programok, azonban a jól induló RealVideo formátumhoz szükséges lejátszóprogram csak kevés készüléktípusban került bevezetésre. Az újabb készülékek felhasználói a 3GP és WMV formátumú fájlokat tudják lejátszani készüléktípustól függően. Az adatbázis videoállományok többek között a kultúra.hu internetes oldalain és a T-zones wapportalon jelennek meg.

Az előbbi példa is bizonyítja, hogy a PORT.hu üzleti megrendelői élen járnak az új technológiai lehetőségekre épülő szolgáltatások indításában, ezért az adatbázis esetében nagyon fontos, hogy frissen bevezetett formátumok használatával is bőséges tartalmat tudjon biztosítani.

A meta-adatok vonatkozásában a PORT.hu megközelítése eltér a könyvtári katalogizálás logikájától. A kép- és videóadatbázis a kulturális programinformációkhoz kapcsolódva (moziműsor, színházműsor, tv-műsor) jelenik meg, önmagában nem kereshető. Ezt azt jelenti, hogy szűk értelemben véve csak néhány meta-adat kapcsolódik hozzá (elnevezés/képaláírás, forrás, készítő, technikai paraméterek).

Legfontosabb lépésként azonban a rögzítéskor a dokumentumot összekapcsolják a PORT.hu teljes adatbázisával: filmekhez, színházi rendezésekhez, tv műsorokhoz, személyekhez vagy helyszínekhez rendelhetik hozzá. Ez egyben azt jelenti, hogy egy film jelenetét tartalmazó fotóhoz tág értelemben számtalan meta-adat kapcsolódik: a film címe, a gyártás éve, a gyártó ország, az alkotók neve és pozíciója, a film ismertetője, a film forgalmazója, a képen szereplő személy neve, születési adatai, életrajza, filmográfiája, valamint a filmhez és a személyekhez kapcsolt további médiaállományok. Ezeknek az adatoknak a figyelembevételével a digitalizált dokumentumokhoz a Dublin Core által javasolt meta-adatok legtöbbször a megfelelője is kapcsolódik.

Az adatbázis könnyen bővíthető újabb kimeneti formátumokkal. Ilyen esetben egy konvertáló eljárást futtatnak le az összes elsődleges példányon és ezzel előállítják az új formátumnak megfelelő másodlagos példányokat. Ezzel egy időben a szerkesztőségi rendszerbe is beillesztik az eljárást, ami így az újonnan feldolgozott dokumentumokból az eddigiek mellett az új formátumú példányt is le fogja gyártani. A közeljövőben várható, hogy a mobiltelefonok növekvő képernyőfelbontásához igazított méretű képméret kerül bevezetésre a WAP oldalakon történő megjelenítésre.

Összefoglalva a PORT.hu kép- és videóarchívuma beváltotta a hozzá fűzött reményeket:

- biztonsággal szolgálja ki a megrendelők és a PORT.hu látogatóinak kéréseit,
- az új dokumentumok hozzáadása a szerkesztőknek minimális idejét veszi igénybe,
- könnyen bővíthető a már feldolgozott dokumentumokra is érvényesen újabb kimeneti formátumokkal.

Az adatbázis korlátait jelenti, melynek átlépése komoly ráfordítással jár:

- ha a tárolt elsődleges példányoknál jobb minőségű kimenetet kell szolgáltatni,
- ha olyan meta-adatra lenne szükség, amely nem elérhető (pl. „képek, amelyeken autó látható”).

Ezek a korlátok a tárolás és feldolgozás költségeinek és a várható igények figyelembevételével lettek megválasztva.

Felhasznált irodalom:

1. PORT.hu médiaadatbázis szerkesztőségi rendszere, felépítése és megjelenési formái. 2005. PORT Computer Rt.
2. dr. Horváth Péter - dr. Koltay Tibor: Digitális könyvtárak és projektek.
<http://www.mek.iif.hu/porta/szint/tarsad/konyvtar/ekonyvt/digkvt/html/index.htm>
3. Prókai Margit: Digitális forma és tartalom.
<http://www.mek.iif.hu/porta/szint/tarsad/konyvtar/ekonyvt/digform/digform.htm>
4. Busznyák János - Dr. Berke József: Képtömörítő eljárások pszichovizuális összehasonlítása laboratóriumi körülmények között – MAMIKA.
http://www.georgikon.hu/digkep/cikk/Busznyak_Berke_cikk_KEPAF4.pdf
5. Image Capture: Hardware and Software.
<http://www.tasi.ac.uk/advice/creating/hwandsw.html>
6. Marcia Stepanek: Data Storage: From Digits to Dust. Business Week, April 20, 1998
7. A Pulman Digital Guidelines magyar változata.
<http://mek.oszk.hu/html/irattar/ajanlas/pulman/index.html>
8. Digital Projects Guidelines (Version 1.3, March 2000), Arizona State Library.
<http://www.dlapr.lib.az.us/digital/index.html>
9. Diane Hillmann: Using Dublin Core. 2003.
<http://www.dublincore.org/documents/usageguide/>
10. Sikeres digitalizálás lépésről lépésre v1.2 Gyakorlati útmutató. 2003. Minerva Working Group 6.
<http://www.nda.hu/Resource.aspx?ResourceID=docstorefile&f=146&t=stored>
11. Samer Alshawwa: Digitizing and organizing information in Libraries, an overview.
<http://rani.sir.arizona.edu/~sirsls/fall03/501/alshawwafp.html>
12. CD-R Media Longevity. <http://www.cd-info.com/CDIC/Technology/CD-R/Media/Longevity.html>
13. Digitalizációs programok a Cseh Köztársaságban. NDA newsletter. 30 April 2004.
http://www.nda.hu/engine.aspx?page=hirlevel_public_02
14. Tom Worthington: Internet-TV Convergence with the Multimedia Home Platform. Communications Research Forum, 26 September 2001.
<http://www.tomw.net.au/2001/itv.html>
15. Knapp Gábor, Lois László, Magyar Gábor, Németh Gergely, Szigeti Szabolcs: A Nemzeti Audiovizuális Archivum műszaki rendszere. 2004.
<http://www.nava.hu/navamuszakirendszere.pdf>
16. Az NDA működése. 2004. http://www.nda.hu/engine.aspx?page=nda_mukodese