

Breinich Gábor

Budapest Főváros Levéltára 10 éves digitalizálási tapasztalatairól

Tisztelt Hölgyeim és Uraim! Tisztelt Kollégák!

A mai napon a levéltári digitalizálásról 13 előadó tart előadást. Jó tíz évvel ezelőtt nagyon nehéz lett volna olyan kollégát találni, aki a témáról érdemben tudott volna nyilatkozni. A készülékek drágák voltak, nehezen hozzáférhetőek, számtalan technikai probléma állt a digitalizálással foglalkozók előtt. A levéltári anyag digitalizálása szakmai kérdésként is igen új volt.

A nemzetközi szakirodalomban 1997-98-ban lehetett az első komolyabb tanulmányokkal találkozni, mint például a Kongresszusi Könyvtár fotóinak digitalizálásáról, a Digitalizálásról, mint megőrzési módról. (Ezeket a tanulmányokat egy a kulturális örökség megőrzésének elősegítésére alakult non profit alapítvány publikálta. (ECPA (European Commission on Preservation and Access 1994-2008 között működött non profit alapítványként a Holland Királyi Művészeti és Tudományos Akadémián)

Mégis 1998-ban egy Budapest Fővárosa által kiírt pályázat segítségével, anyagi lehetőség nyílt arra, hogy levéltárunk megindíthassa digitalizálási programját.

Az első év az eszközök kiválasztásával, a lehetséges technológiák tanulmányozásával telt, miközben kialakítottuk az alapelveinket.

Azt tűztük ki célul, hogy az elsősorban a képi információkat őrző anyagok körében tegye lehetővé azt, hogy a gyakran használt dokumentumok használatát, amelyek állagukban gyorsan romlanak, és a sérülések miatt fennáll az információ vesztes veszélye, olyan minőségű digitális másolattal váltsuk ki, amely megőrzi minden leolvasható információs elemet, számítógépen is lehetővé teszi a kutatást, és az eredetivel azonos minőségű másolat „hard copy” készítését is biztosítja.

Sok kísérlettel teszteltük a különböző berendezéseket, amelynek során kiválasztásra került egy A/0 méretű roll szkennerek, és a hozzátartozó perifériák, a nyomtatásra alkalmas plotterig bezárólag. A tesztelésnél a legnagyobb adatsűrűséget mutató térképünkkel kísérleteztük ki azt a felbontási küszöbértéket, a 400 DPI optikai felbontást, amelynél minden digitalizált információ tisztán, jól olvashatóan, nagyíthatóan megjelent. A megőrzést optikai lemezekre végeztük, bízva abban, hogy azok a hosszútávú megőrzés céljaira alkalmasak lesznek.

Úgy döntöttünk, hogy adatvesztés nélküli képfórmátumban (TIFF) őrizzük az archiválásra szánt példányokat, és adatvesztéses tömörítési eljárással (JPG) készítjük el a kutatói használatra szánt változatokat.

A nagy tárolási méret ebben az időben még komoly technikai problémát jelentett, (Erről ma már hajlamosak vagyunk elfeledkezni a terrabájtok korszakában.), ezért reális kompromisszumként az archív állományok esetében a színárnyalatok számát csökkentettük le 256-ra, ami az 5-6 szint használó térképek és tervrajzok esetében nem okozott információvesztést.

A saját műhelyben, egy fővel működő digitalizáló „részleg” évi 1000 térkép feldolgozását végezte el, az igen gyakori egyéb kérések, kutatói és intézményi kérések teljesítése mellett.

Az első évek pozitív tapasztalatait a következőkben lehetne összegezni:

- A digitalizálási programban a feldolgozandó anyag kiválasztása jól történt meg.
- A kiválasztott térképgyűjtemény olyan értékes és intenzíven használt irat együttes, amelynek feldolgozását belátható időn belül be is lehetett fejezni.

- A digitalizálás állományvédelmi szempontból biztonságos módon történt
- . Az elkészült digitális állományok biztonságos tárolása megoldott volt, a kutatók a levéltáron belül használhatták a rendelkezésükre álló állományokat, és ami a legfontosabb, megfelelő segédletek álltak rendelkezésre a kutatáshoz.
- A költségek jól kézben tarthatóak voltak, lényegében saját erőket használtunk, kevés anyagfelhasználással.

A negatív tapasztalatok a következők voltak:

- Az A/0 méretnél lényegesen nagyobb térképekkel nem tudtunk mit kezdeni.
- Gond volt a vastag, igen gyakran már eldeformálódott kartonokra, vagy falapokra felkasírozott térképek kezelésével. A gördítő szerkezet vagy nem tudta befogadni a túl nagy vastagság miatt, vagy pedig egyenetlenül továbbította a dokumentumot, így torzult a felvétel.
- A teljes munkafolyamat az egy munkaállomásos rendszerben egy emberre hárult, (szkennelés, mentés, adatkonverzió, CD lemezek megírása, nyilvántartás készítése), ami erősen csökkentette a hatékonyságot.

Az észlelt problémák megoldására külső erőforrásokat vettünk igénybe. A nagyméretű, vagy vastag hordozóanyagra kasírozott térképek digitalizálását egy fotósra bíztuk. A fotós, aki az erre a célra konstruált fekvő, és sínen mozgatható tárgyasztala fölé, 5 méter magasan elhelyezett digitális kamerájával pásztázta végig több felvétellel a nagyméretű térképeket, utóbb számítógépén összeillesztette a képeket.

2003-ban a levéltárunkban őrzött, kiemelkedő értékkel bíró 19. századi fotóalbumok (Klősz György fotográfus anyaga) digitalizálására egy pályázat adott lehetőséget. Az albumokban lévő 2026 kép digitalizálását, azok szétbontása nélkül egy erre a célra alkalmas, saját konstrukciójú reprodukciós asztalon, szintén külső vállalkozó végezte el.

Mindkét munka alkalmával a digitalizálásra meghatározott alapparaméterek, valamint az állományvédelmi előírások betartása kötelező feltétel volt.

Ez alkalommal azt a célt is kitűztük, hogy az Interneten keresztül is lehessen a fotóanyagot kutatni. Az erre a célra történt szoftverfejlesztés segítségével 2004 óta az elkészült anyag adatbázisa és képei a levéltárunk honlapján online kutathatóvá váltak.

2003-ban indult el a hangzó anyagok digitalizálása, ahol leggyakrabban a testületi üléseken felvett, és magnetofonszalagokon rögzített anyagok új adathordozóra mentését, zörejektől való szűrését, a hallhatóság javítását végezzük el az optikai adathordozóra mentés előtt. Itt is a szabványosnak tekinthető formátumokat használjuk (Wave, mint nem tömörített és MP3 tömörítést.)

Öt év elteltével az informatikai technika világa jelentősen megváltozott, és megjelentek azok a vállalkozások is, amelyek a levéltári információs vagyont feldolgozásában már üzletet láttak. Egyrészt a pályázatokkal támogatott digitalizálás jelentett üzletet, másrészt pedig a digitális kiadványok kiadása terén, magának az információs vagyont értékesítési lehetősége is felkeltette a vállalkozók érdeklődését.

Ennek jóvoltából felgyorsulhatott a digitalizálás tempója. A térképsorozatunk digitalizálásába beszállt vállalkozó közreműködésével, aki hozzánk telepítette gépeit és embereit, ellenőrzött körülmények között, egy éven belül befejezhattük a teljes térképállományunk feldolgozását. Ez levéltárunk és a vállalkozó közös kiadásában 2005-2006 között négy DVD lemezen katalógus adatokkal együtt jelent meg. Digitalizálásra, és DVD-én kiadásra került Pest város 1861-1873 közötti építészeti tervanyaga is.

A sikeren felbuzdulva 2007-ben ugyanebben az együttműködésben sikerült megjelentetni Buda és Pest történeti topográfiáját, amelyik egy jelentős újdonságot is hozott. A térinformatika nagyot lépett előre, és sikeresen oldotta meg a korábbi időszakokban keletkezett térképek mai térképvetületbe illesztését, amely eljárást georeferálásnak neveznek. Ennek segítségével megoldhatóvá vált egy adott pont különböző időkből, különböző térképeken való megtalálása, a változások egy képernyőn való összevethetősége, távolságok pontos lemérése az egyes térképeken, ami nagyszerű felhasználási lehetőségeket nyújt a kutatók számára.

A külső erőforrások bevonása, jelentősen felgyorsította a digitalizálási munka tempóját és kiterjedtségét is. Az automata dokumentumszkennerek ezerszámra nyelik az oldalakat és a jó minőségű nyomtatott, vagy írógéppel írt szövegek automatikus felismertetése szintén hatalmas ugrást jelent a digitalizálás és közzététel szempontjából.

Mára már az Interneten olvashatók szöveges források közül Budapest Közgazdasági és Munkaadók Szövetségének jegyzőkönyvei az 1873-1949 közötti időszakból, az MDP/MSZMP budapesti vezető testületeinek jegyzőkönyvei, mégpedig a karakterfelismertetés jóvoltából teljes szövegűen kereshetően, azonban a PDF technika segítségével az eredeti megjelenési formáját megőrizve, hitelesen jelenítve meg az irat hű mását. Digitalizálás után az Internetre kerültek továbbá a levéltárunk korábban nyomtatásban megjelent kiadványaiból is a fontosabbnak ítélt művek. (Text PDF, avagy kétrétegű PDF technika segítségével.)

Ezekre a vállalkozói erőforrások igénybevitelével kerülhetett sor, ahol az eszköz és élőmunka kihasználtsága lényegesen rugalmasabban szervezhető meg, mint egy közgyűjteményen belül.

A digitalizálás egyre inkább a napi gyakorlat része, és beépül az állományvédelmi gyakorlatban a mikrofilmlező kamera mellé, olykor helyébe. A mikrofilm digitalizálás megindítása 2011-ben várható levéltárunkban. 15 millió felvétel feldolgozását tervezzük az elkövetkezendő években. Éppen a napokban kaptuk meg anyakönyvi mikrofilmjeink 1.2 terrabájtnyi digitalizált anyagát a mormon egyházzal való együttműködésünk keretében.

Néhány adat az elmondottakkal kapcsolatban:

- 1999-2003 között 5000 térkép digitalizálását végeztük saját erőből.
- 2005-2006-ban külsős közreműködővel 9000 felvétel készült el, és fejeződött be a térképanyag feldolgozása.
- Az építészeti tervanyagból kb további 27.000 felvétel készült el. Itt már az adatbázis készítés lemaradása hátráltatja csak a közzétételt.
- Jelenleg a BFL 3 terrabájtnyi digitalizált állománnyal rendelkeznek, és ez felveti a megfelelő tárolás kérdését is.

Az optikai lemezeket tárolt anyagainkat, ebben az évben megkezdjük egy RAID 6 biztonsági rendszerrel kialakított mágneses adattárolóba betölteni, ahol az őrzés, és kezelés is jobb színvonalú, mint a CD és DVD lemezeken való őrzés. Az optikai lemezek megmaradnak, biztonsági példányként.

A végleges megoldás, a hosszú távú archiválást kérdését 2011-ben megoldó elektronikus levéltár katasztrófatűrő rendszerében való letárolás lesz.

Az 1998-ban kialakított alapelvek az elmúlt időszakban kiállták a próbát. Nem kellett újra feldolgoztatnunk anyagokat rossz döntések, vagy hibák miatt.

A választott felbontási szint, és az adatformátumok az archív állományok esetében időtállóknak bizonyultak, és a felhasználók számára készített különböző formátumokba átkonvertálhatók voltak.

A digitalizálási programok előtt időben megkezdett, ill. párhuzamosan folytatott adatbázis építési folyamat lehetővé tette a digitalizált anyag közzétételét, használatát.

A saját műhely feladatköre részben átalakult, az aktuális kutatói igények, valamint a levéltár kiállítási célú igényeinek megnövekedésével, de úgy tűnik, az időszakonként lökésszerűen jelentkező igények kielégítésére a vállalkozói szféra hatékonyan (és költséghatékonyan!) vonható be, és jól együtt tud működni, a levéltári iratanyag közelében, megfelelő szakemberekkel rendelkező saját levéltári digitalizáló műhellyel.

A magyar levéltári iratanyag digitalizálásával kapcsolatos eredményeket a Magyar Levéltári Portálon lehet áttekinteni a legkönnyebben. A címe www.archivportal.hu

Köszönöm megtisztelő figyelmüket!