

VALÓSÁG

Kerekasztal-beszélgetés a digitalizációról, az emlékmentésről és az ezzel együtt járó szelekcióról

A beszélgetés résztvevői: *Biszak Sándor* – cégvezető, Arcanum Adatbázis Kft.; *Gárdos Judit* – kutató, MTA Társadalomtudományi Kutatóközpont Szociológiai Intézetének Kutatási Dokumentációs Központja, valamint a 20. Század Hangja Archívum és Kutatóműhely; *Ládi László* – osztályvezető, Az Országos Széchényi Könyvtár Könyvtárintézet Kutatási Osztály; *Mikó Zsuzsanna* – főigazgató, Magyar Országos Levéltár; *Nagy Péter Tibor* – egyetemi tanár, ELTE TÁTK OITK/Wesley Egyház- és Vallásszociológiai Kutatóközpont.

Educatio: A digitalizálás, emlékmentés témájához kapcsolódóan e kerekasztal-beszélgetés során ennek hazai környezeti feltételeivel, szervezeti és jogi kereteivel, szereplőivel és érdekcsoportjaival foglalkoznánk, és persze beszélünk arról, hogy hogyan írható le és milyen következménye van a mentéssel óhatatlanul együtt járó szelekciónak. Mindenekelőtt azonban bemutatkozásként kinek-kinek a digitalizálással kapcsolatos tevékenységi körét vennénk sorra. *Mikó Zsuzsanna*: A Magyar Nemzeti Levéltár főigazgatójaként a levéltárak iratainak hozzáférhetővé tétele érdekében végzett egyre hangsúlyosabbá váló tevékenységünk révén kapcsolódni a beszélgetés témájához. Esetünkben a digitalizálás funkciója elsősorban az, hogy minél szélesebb körben hozzáférhetővé tegyük azokat az információkat, amelyek korábban csak a levéltár kutatótermeiben voltak elérhetőek. Ma már egyre több anyagot publikálunk honlapunkon is, emellett igyekszünk lehetővé tenni, hogy a kutatótermeinkben dolgozó kutatók már digitális formában nézhessék meg az anyagokat. A digitalizálás nagyon széles körre kiterjed. A papíralapú anyagok digitalizálása mellett hanganyagok, mozgóképi anyagok, mikrofilmek digitalizálásáról is beszélhetünk. Kicsit a nagyságrendekről: a Magyar Nemzeti Levéltár főintézményében, az Országos Levéltárban jelenleg nyolcvanhárom kilométernyi iratanyag van, elég nagy tehát a kihívás, ha ezeket hozzáférhetővé kívánjuk tenni. Ebben egyébként európai összehasonlításban is nagyon jól állunk, hiszen az Országos Levéltár kb. tízmillió digitalizált tartalomnál tart, amelyek honlapon keresztül hozzáférhetővé is tehetőek. Fontos azonban látnunk azt is, hogy a digitalizálásnál – melynek technológiája mára meglehetősen egyszerű, gyors és olcsó – jóval nagyobb kihívás az, hogy a digitalizált anyagok tartalmához, kereshetőségéhez biztosítsunk megfelelő adatbázisokat. Ezzel kevésbé állunk jól, mint a digitalizálással.

Gárdos Judit: A Magyar Tudományos Akadémia Szociológiai Intézetében már évek óta – immár egy második OTKA kutatás bázisán – működik Kovács Éva vezetésével a 20. Század Hangja Archívum és Kutatóműhely.¹ Alapvetően az elmúlt ötven év társadalomtudományi, szociológiai kincseit szeretnénk megőrizni. Főként interjú kutatási anyagokra koncentrálunk, ezek hanganyagaira és leirataira, valamint az ezekhez kapcsolódó képekre, videókra. A papír alapon feldolgozott anyagokat az Open Society Archives

1 A 20. Század Hangja.

(OSA) együttműködésével tároljuk, és a digitalizált tartalmat az OSA rendszerében tesz-
szük hozzáférhetővé egy DSpace nevű ingyenes Open Source Repository rendszerben.²
Maga a digitalizálás valóban viszonylag egyszerű része a folyamatnak, a meta-adatolás a
komolyabb kihívás. Ennek színvonalas szakmai megvalósítása nagy befektetést igényel,
márpedig jelenleg erre Magyarországon nem áldoznak igazán szívesen. A szelekció ne-
héz kérdés. Mi alapvetően nagy klasszikus szociológiai gyűjtemények anyagait dolgoz-
zuk fel. Losonczi Ágnes gyűjteménye, H. Sas Judit gyermekrajz-kutatásának anyagai, az
MTA Szociológiai Kutató Intézetében a 70-es évek közepén készült munkáséletmód in-
terjúk csaknem tizenhatezer oldalnyi anyaga és még sorolhatnám. Egyfelől tehát a men-
tésre koncentrálunk, olyan anyagok digitalizálására, amelyek máskülönben elvesznének.
Másfelől pedig vannak a közelmúltban, hazai és nemzetközi kutatások keretében létre-
hozott gyűjteményeink is, amelyek eleve digitálisan készültek. Ilyen pl. a Kovács Éva ál-
tal vezetett, az 1939–1945 közötti munkaszolgálatot feltáró interjú-és fotógyűjtemény,
csak hogy egyet említsék.

Ládi László: A könyvtárak mindig is a nyomtatott dokumentumok legfőbb megőrző helyei
voltak. Az elektronikus dokumentumok térhódításával ezeknek is helyet kell biztosítani a
könyvtári keretek között. Ezek gyűjtését, bekerülését egyelőre kétségessé teszi, hogy a ki-
adók egyre kevésbé akarják az elektronikus könyveket kötelempéldányként beszállítani.
Erre megvan a maguk természetes indoka, nem érzik ugyanis biztonságosnak az elektroni-
kus tartalom felhasználását. A digitalizálással valamilyen szinten a könyvtárak a technol-
gia megjelenésének kezdetétől foglalkoztak, például a legtöbb megyei könyvtárban zajlottak
helyismereti digitalizálások, de az egyetemi könyvtárak is építenek nagy repozitóriumokat.
Mondhatjuk, hogy az országos könyvtárak eszközszinten gyakorlatilag készen állnak a di-
gitalizálásra és sok mindent digitalizálnak is. Elsősorban állományvédelmi digitalizálás
folyik, azaz a nagyon rossz állapotban lévő folyóiratok és egyéb nyomtatványok megőrzése.
Biszak Sándor: Az Arcanum Adatbázis Kiadó digitalizálással foglalkozik, elsősorban a
második világháború előtti anyagokkal. Nagyon sokat dolgozunk levéltárakkal, könyv-
tárakkal. Külön projekt keretében digitalizálunk tudományos folyóiratokat is, ezidáig
százhusz, százharminc folyóirat teljes állományát digitalizáltuk, olyanokét, mint az At-
henaicum, a Vasárnapi Újság vagy a Századok. A teljesség igényével dolgozunk, de persze
azt megmondani, hogy a Századoknak hány melléklete volt, ez hány oldalon jelent meg
és hol található, még az Országos Széchényi Könyvtár szolgálati katalógusának felhasz-
nálásával is igen nehéz. Jellemzően nyolc, tíz könyvtárból tudunk összeszedni egy teljes
sorozatot egy folyóiratból, melyeket oldal szinten dolgozunk fel és ellenőrünk. Havonta
fél millió oldalt digitalizálunk, s én azt érzem inkább, ahogy fogynak el az anyagok. Ki
gondolta volna még 1989-ben, a cég alapításakor, hogy 2010-re a magyar szabadalmak
teljes állománya digitalizált és kereshető módon hozzáférhető lesz? Nem nagyon találunk
olyan jelentősebb folyóiratot, ami ne lenne digitalizálva. Jelenleg a napilapokkal dolgo-
zunk (Népszava, Pesti Hírlap, Pesti Napló), s bár ez tényleg jelentős mennyiség, az már
látható, hogy jövő ilyenkorra elfogynak ezek is. Fogytán vannak az értelmes anyagok. S
míg a mi kapacitásainkkal egy hónap alatt félmillió oldalt digitalizálunk, addig a Goog-
le ugyanannyi idő alatt félmillió könyvet.

Nagy Péter Tibor: A Wesley Egyház- és Vallásszociológiai Kutatóközpont keretein belül
– részben önkéntesek bevonásával – több olyan projekt is zajlik, amelyek talán példaér-

² OSA Archivum.

tékűek lehetnek abból a szempontból, hogy hogyan lehet – komolyabb költségvetés nélkül – a legnagyobb mértékben érvényesíteni a tartalom-feltárás lehetőségeit. Az egyik nagy és széles körben ismert projektünk a Magyar Zsidó Lexikon digitalizációja, amely radikálisan különbözik egy „Google-típusú” pusztá szkennelési tevékenységtől. Itt több ezer órányi munkával készültek mutatók a digitalizált tartalomhoz. Én ezt a digitalizációs fordulat nagyon fontos tartalmi elemének tekintem, hiszen így egy köztes állapot jön létre a hagyományos katalógusok és a majdnem reménytelen full textes keresés között. Másik nagy projektünk, az 1929–31-es négykötetes Magyar Színművészeti Lexikon ugyancsak erős tartalmi mutatókkal rendelkezik. Különösen büszke vagyok arra – és ez tényleg szigorú szerzői jogi kérdés –, hogy Randolph L. Braham professzor úr rövid beszélgetés után átlátta a jelentőségét, hozzájárult és örült neki, hogy jelentős művét, a „Magyar Holocaust”-ot digitalizáljuk és adatbázisba szervezzük. (Minderre rá lehet keresni a Magyar Elektronikus Könyvtárban.)

E: Milyen korlátokat jelentenek a digitalizálásban a jelenleg érvényes hazai szerzői jogok?

L. L.: Az Országos Széchényi Könyvtárban egy és kétmillió közötti digitalizált dokumentum és tartalom van, ám az ezekhez való hozzáférés kérdéses. Be kell tartani a jogszabályokat és törvényeket, amelyek általában úgy száznegyven évre visszamenőleg nem nagyon teszik lehetővé a legális digitalizálást. A jelenleg zajló, országos szintű elektronikus dokumentumszolgáltatást célzó fejlesztés legnagyobb problémáját az jelenti, hogy a szerzői jog olyan korlátokat állít elő, amelyet nagyon nehéz megkerülni. Például, ha valaki 1870-ben, húsz évesen írt egy könyvet és mondjuk kilencven éves kort ért meg, akkor az 1940-es években bekövetkezett halála után még hetven évig tart a műre a szerzői jogvédelem. Illetve, ha későbbi szerzőről van szó, akinek azonban számos örököse van, ezek mindegyikét meg kell keresni ahhoz, hogy a művet a szerzői jog szabályai szerint hozzáférhetővé tehesük. S ragozhatnánk tovább is, hisz egy folyóirat valamennyi cikkénél meg kellene keresni a szerzőt, sőt, ha már megszűnt folyóiratról van szó, akkor az eredeti kiadói jog birtoklóit is. Ez okozza tehát a legnagyobb problémát a legális könyvtári digitalizálásban. Miközben mindannyian tudjuk, hogy tízezer számra keringenek magyar nyelvű könyvek az interneten, az egész frissen megjelent köteteket is beleértve. Illegálisan tehát hozzáférhetőek, ám a könyvtárak legálisan ezt nem tudják biztosítani.

N. P. T.: Azt gondolnám, hogy a még élő szerzők, kutatók az esetek kilencven százalékában a művük digitális megjelentetésére vonatkozó felkérésre igent mondanak, hiszen ez az érdekük is. Természetesen sok ezer szerzős napilapoknál ezt végigvinni lehetetlenség. De ilyen esetekre léteznek megkerülő lehetőségek is. Ilyen például a kutatási célú digitalizálás. Egy oktatási intézmény kizárólag kutatási célra elvégezheti a digitalizációt, a tartalom közzétételében a szerzői jog pedig már az általa, saját szellemi munkájának hozzáadásával létrehozott műre vonatkozik. Mentsvár lehet ez akkor, ha például az eredeti dokumentumnak sok szerzője van, vagy a szerzők köre meghatározhatatlan.

L. L.: Megjegyezném, hogy a felmérések mind azt mutatják, hogy annak a szerzőnek, aki szabadon hozzáférhetővé teszi a szépirodalmi műveit, a hagyományos forgalma megnő. A digitális hozzáférhetőség lehetősége növeli az ismertséget, kihat a vásárlásokra.

N. P. T.: Itt van a torrent-világ. A hatalmas erőforrásokat megmozgató kiadók, filmstúdiók sem tudják megakadályozni, hogy alkotásaik tömegesen fölkerüljenek a torrentre, nyilván tehát a szerzők és örököseik sem lehetnek ebben igazán hatékonyak. Ez a dolog átszakadt, tudomásul kell venni a digitális tartalmak visszafordíthatatlan elterjedését. S

én azt gondolom, ez a dolgok normális rendje. Az első írott szövegek óta több ezer év telt el szerzői jog nélkül, fél évezrede van nyomtatott könyv is, elképzelhető tehát, hogy a szerzői jog két évszázada csak egy átmeneti időszak az emberiség életében.

L. L.: A szerzői jog kérdésében magam is a liberálisabb vonalat képviselném. Fontosnak tartom a folyamatos digitalizálást, mégpedig a legfrissebb műveket, hiszen ezekre van a legnagyobb igény. Természetesen megszüntetjük a nyilvánosság tételét, amennyiben a szerző ezt nehezményezi, ám tudatjuk vele azt is, hogy ezzel kétféle potenciális felhasználó elől zárja el a művét, radikálisan csökkentve annak ismertségét. Látnunk kell azt is, hogy miután a könyvtárak megvették ezeket a könyveket, azok a polcokról elvihetők, a lapok fénymásolhatók, majd terjeszthetők függetlenül attól, hogy a könyvtár ezt digitalizálta-e vagy sem. Ráadásul a szerzői jogi törvény nagyon egyszerűen megkerülhető. Az anyag egy része ugyanis digitalizálható, mely egy második digitalizálási szakaszban kiegészülhet a hiányzó résszel. Kérdés az is, hogy mi egy része a műnek? Végh Antal futballról szóló könyvét például egy másik kiadó az első három oldal, a bevezető és a címlap nélkül adta ki, mondván, hogy ez a mű egy része csupán. A bíróság ezt elfogadta.

E: Miben más a kutatási adatbázisok, levéltári anyagok hozzáférhetővé tétele?

G. J.: Fontos megkülönböztetnünk a különböző jellegű dokumentumokat a szerzői jog szempontjából. Az irodalmi művek szerzője vagy annak örökösei anyagilag érdekeltek a könyv eladásában, nem feltétlenül érdekük tehát a digitális hozzáférés. A levéltári anyagok esetében ugyanakkor a szerzői jog kérdése nem releváns. A kutatási anyagok szerzőjének kérdése megint csak komplikált dolog: például egy interjú esetében a kérdező vagy a válaszadó minősül szerzőnek? Erről a problémáról még évekkor ezelőtt hosszasan konzultáltunk az ombudsmannal. Akkor az lett a konklúzió, a kutatási céllal készült dokumentumok kutatási célra a későbbiekben is felhasználhatóak. Kivéve persze, ha másfajta megállapodást kötött a válaszadó a kutatókkal.

M. Zs.: A Magyar Nemzeti Levéltár egy közhiteles intézmény. Teljesen mindegy, történetesen egyetértünk-e a hatályos jogszabályokkal vagy sem, egy ilyen intézmény nem engedheti meg magának, hogy a számára átadott anyagok kezelése kapcsán jogszabályokat sértsen meg. Ez vonatkozik a gyakran emlegetett egyházi anyakönyvekre is, amelyek bekerültek a Levéltárba, miután ezeket a mormon egyház az ötvenes évek végén, hatvanas évek elején belügyi támogatással mikrofilmen rögzítette. Ezek az egyházi anyakönyvek nem a miénk, hanem az egyházaké. Milyen jogon tehetnénk közzé az interneten? Sőt, bár nálunk vannak, megvannak, a kutatók számára sem adhatjuk ki, hiszen ismétlem, nem a miénk.

B. S.: Megdöbbenő ehhez képest a francia gyakorlat, ahol a több tucatnyi regionális levéltár különböző szoftverekkel működő egyházi anyakönyvi adatbázisai mind hozzáférhetőek, beleértve az egykori gyarmatok (mint például Mozambik) anyakönyvi adatait is. Itt is azt teszik közzé, amit a mormonok annak idején mikrofilmen rögzítettek, és a későbbiekben a levéltárak digitalizáltak. Ilyen jellegű és nagyságrendű együttműködésről nem is tudok ezen kívül.

N. P. T.: Emlékeztetnék arra, hogy Mária Terézia az egykori jezsuita vagyion hasznosítását célzó rendeletének értelmezésekor a magyar parlament 1870-ben úgy döntött, hogy az egyházi célra tett adományok, minthogy olyan egyházi funkcióval voltak kapcsolatosak, melyek később közfunkciók lettek, közvagyonnak számítanak. Ilyen értelemben tehát – jogértelmezésem szerint – az anyakönyvi adatok közvagyonnak számítanak, mellyel a Magyar Nemzeti Levéltár rendelkezhet. Megérne egy próbapert. Abban mintha többé-ke-

vésbé egyetérténénk, hogy ezek a jogszabályok – felhasználóként, kutatóként, szolgáltatóként – leginkább akadályoznak minket. Akadályozzák magát a tudományos kutatást is. A jogszabály azonban nem valamiféle szent szöveg, hanem egy társadalmi valóság, amely (úgy szövegét, mint későbbi értelmezését tekintve) különböző érdekcsoportok küzdelméből alakul ki. Ilyen értelemben ezen érdekcsoportok – könyvtárak, levéltárak, kutatói szervezetek – képviselőinek jelentős lobbizási tevékenységet kell folytatniuk a változtatás érdekében. Illetve adódhat az a megoldás, hogy mindenki szabadon digitalizálja a szükséges tartalmakat, vállalva az utólagos pereskedést.

G. J.: Ez a kutatási adatok esetében nem működhet. Az eleve nyilvánosságnak készült könyvek, folyóiratok esetében mondhatjuk, hogy kerüljön minden nyilvánosságra. A kutatási adatok esetében azonban egy ilyen hozzáállással a kutatás alapvető bizalmi bázisát ássuk alá. Rettenetesen sok adattisztítási munka húzódik meg a kutatási adatbázisok anonimizálása mögött, hiszen arra is tekintettel kell lennünk, hogy több adatbázis integrálása esetén se lehessen identifikálni az egyént, akinek ebből történetesen kára is származhat. És itt nem az esetleges pereskedésről van szó, hanem arról a bizalmi helyzetről, hogy a válaszadók önként segítenek nekünk azzal, hogy a kutatásban részt vesznek. Ebből nekik káruk, problémájuk később sem származhat. Ezek az adatok adott esetben nagyon szenzitívek lehetnek (szexuális orientáció, betegség stb.).

N. P. T.: A történettudományi kutatásokban évtizedek óta kialakult egy norma, hogy hol, hány év után teszik kutathatóvá a különböző adatokat, akár személynevekkel is. Ez még akár a titkosszolgálati adatokra is kiterjed. S ez akkor is így van, ha ez az adatbázisban szereplőnek történetesen sérti az érdekeit.

G. J.: Adminisztratív adatbázisokban szereplő adatok e tekintetben nem hasonlíthatóak össze kutatási adatokkal. Hiába veszik fel névtelenül a kutatási adatokat, azok kombinációja révén az egyén azonosítható lehet. Például a lakóhely és foglalkozás közös halmazán nem nehéz azonosítani a falusi papot. A nemzetközi kutatások esetében is nagy befektetett munkával végzik el az anonimizálást. A 20. század Hangja Archivum és Kutatóműhelyben követett eljárás szerint mi nem anonimizálunk, hanem az adatkérővel kutatási nyilatkozatot íratunk alá, abban bízva, hogy az anyagban lévő információkat név szerint nem használják fel. A külföldi gyakorlatban rendszerint az interjúkat is csak anonimizálás után adják ki még a kutatóknak is.

E: *A tartalmak nyilvánosságra hozatala óhatatlanul rendszerezést is jelent. Milyen elv lehet ezek mögött a struktúrák mögött?*

G. J.: Hiába digitalizálunk mindent, ami csak létezik a világon. Hozzáadott érték, alapesetben megfelelő hozzáférési felület nélkül a tartalom csak porosodni fog egy szerveren.

B. S.: Ebben a Google verhetetlen.

L. L.: De mit kezdünk a Google-lal, amikor a „Budapest” szóra sokmillió találatot ad ki? Könyvtárosi megközelítésben ez használhatatlan.

N. P. T.: A nemzeti könyvtárak egyik alapvető funkciója, hogy adatbázis alapon digitálisan kereshetővé tegyék a nemzeti könyvtárlományt. Erre szolgál az OSZK keresőrendszere, újabb folyóiratok esetében – s gyakran visszamenőlegesen is – a MATARKA, más könyvek kapcsán a MOKKA. Ezt a fajta adatbázis-építési tevékenységet tekinthetjük a hozzáférés biztosítása nulladik körének, amelynek ellátása továbbra is a könyvtáros szakma feladata marad. Ezt a hatalmas tartalomfeltáró munkát már csak azért sem végezheti el más, mint a nemzeti könyvtárak, mert a világon sok száz nyelven nyomtatnak kiadvá-

nyokat, és sok száz különböző kultúrkör valamifajta ismerete kell ahhoz, hogy mindezt tisztességesen tárgyszavazza valaki. Ez az idők végezetéig eltart. Ugyanez igaz véleményem szerint a levéltári iratokra is. Tudomásul kell venni, ahogy egyre több dolog kerül fel az internetre, egyre több értelmes találatot generálva, az interneten nem elérhető dolgokat egyre kevésbé fogják olvasni. A Magyar Zsidó Lexikon digitalizálási munkálatai során az egyik legfontosabb motívumom az volt, hogy megmentssem azt az óriási szellemi munkát, amit az elmúlt kétszáz évben a legkülönbözőbb lexikonszerkesztők, korrektorok elvégeztek. Ezt a munkát nem lehet kiváltani azzal, hogy majd a Wikipédiára vagy más újonnan létesített site-okra így vagy úgy feltöltődik valamiféle tartalom. Számomra ez a kulcselem abban, hogy nem rendszerezetlen internetes adatközléseket kívánunk szaporítani, hanem az évtizedekkel ezelőtt megjelent kézikönyveket vagy folyóiratokat teszünk nyilvánossá: ezzel ugyanis ténylegesen többet mentünk meg, mint maga a szöveg. Évszázadok szerkesztői és ellenőrzési munkáját.

M. Zs.: A levéltári adatok hozzáféréseben is értelmezhető a rendszerezés, szelektálás szerepe. A Levéltár adatállományának jelentős része a honlapon keresztül mindenféle regisztráció nélkül, pusztán keresés alapján hozzáférhető. Vannak emellett olyan adatok, melyeket csak speciális megrendelést követően, kizárólag tudományos kutatók számára teszünk hozzáférhetővé, akár online formában is. Nyilván a weblapunkon olyan adatokat igyekszünk publikálni, amelyekről azt gondoljuk, hogy sokakat érdekelnek. Akár MSZMP jegyzőkönyvek, oklevelek vagy összeírások.

G. J.: A rendszerezés munkája – a digitalizálással ellentétben – soha véget nem érő. Például a régi térképek digitalizálása után ezek adatait összeilleszthetik népszámlálási adatokkal, amiből új adatbázisok jönnek létre. Miután az összes térképet digitalizáltuk, a világ valójában afelé fog mozdulni, hogy ezen bázison milyen minőségi tartalom és hozzáadott érték születik.

B. S.: Ez már a tudományos kutatás feladata. Magyarországon az összes régi térkép – 250–300 ezres nagyságrend – digitalizálva van már. Ahol a digitalizálásnak vége, ott kezdődik a feldolgozás, a kutatás. Mi ehhez biztosítjuk a forrásokat. A mi digitalizálási munkánkat egyébként a Google hatalmas kapacitásain kívül a közösség növekvő szerepe is behatárolja. Ezzel együtt, akár csak a Google-t, a Wikipédiát is a világ egyik nagy találmányának tartom, mint önkéntesen együttműködő – azt mondanám, okos – emberek szövetségét.

M. Zs.: Van levéltári gyakorlat a közösség bevonására a tartalmak digitalizálásába. Ennek során digitális tartalmakat teszünk hozzáférhetővé olyan közösségek számára, akik hajlandóak azokat feldolgozni. Az általunk kifejlesztett, elkészített adatbázis sémába az önkéntesek saját otthoni gépeikről rögzítik az adatokat. Egy ilyen forrásból kapott információt akkor fogadunk el hitelesnek, ha három egyező adatot kapunk. Hatalmas mennyiségű olyan alapanyag van, amelynek feldolgozása csak ebben az önkéntes formában valósulhat meg.

B. S.: Tudtommal ilyen az 1828-as összeírás, amely háztartásonkénti bontásban tartalmazza az összes magyar adózó adatait. Maga az adat digitalizálva van, többszázezer oldalnyi terjedelemben, de az ezekhez tartozó négymillió nevet kézzel be kell vinni valahogy. Nagy valószínűséggel se állami, se uniós forrás nem áll majd rendelkezésre a négymillió név beírásához, így tehát a lokális közösség, az adott falu, város vagy kerület önkéntesei jelenthetik az egyetlen esélyt az adatok megmentésére. Ez a mormonok modellje is, akik az egyházi

anyakönyvek feldolgozása során szintén önkéntesekkel dolgoznak, és a Magyar Család-kutatók Egyesülete is ugyanezt követi a polgári házassági anyakönyvek feldolgozása során. *M. Zs.:* Ez valóban fontos erőforrás, amely a maga hibázási lehetőségeivel együtt sem nélkülözhető. Ennek mérlegeléséhez látnunk kell egyrészt, hogy a levéltáros is ember, aki tévedhet, és hogy a hibák többszörös ellenőrzéssel minimalizálhatók. Másrészt döntenünk kell, hogy vagy vállaljuk a feldolgozásnak e kevésbé professzionális módját, vagy együtt élünk azzal a ténnyel, hogy az említett adatok még az unokáink idején sem lesznek feldolgozva.

E: A digitalizálás biztos és hatékony módja a dokumentumok megmentésének?

L. L.: A digitális tartalmaknál figyelembe kell vennünk, hogy hihetetlen gyorsasággal avulnak el. Egy húsz évvel ezelőtti adathordozó már aligha használható, folyamatos karbantartást igényel. Ugyanez vonatkozik a megjelenítő eszközökre.

M. Zs.: Ennek a problémának a kezelése a közgyűjtemények fontos feladata, és mondhatjuk, hogy a Levéltár ebben előre jár. Nagyon komoly erőforrásokat fordítunk arra, hogy az elektronikus iratot, digitális tartalmat hosszú távra megőrizzük. Sajtótájékoztatókon szoktam mondani, hogy ha a legkorábbi iratunk 1109-ből származik, akkor a mostaniak esetében is minimálisan ez az időtáv legyen az irányadó. Nem is halaszthatjuk a szembe-sülést ezzel a problémával, hiszen ma már az iratok jelentős része csak elektronikusan keletkezik, ezeket értelmetlen volna papír alapon is őrizni. Ez a közgyűjteményekre olyan értelemben is feladatot ró, hogy nekik kell a hosszú távú megőrzéshez szükséges anyagi forrásokat is biztosítani. Az bizonyos, hogy a Google arra nem fog erőforrást áldozni, hogy ezek a tartalmak nyolcvan év múlva is biztonságosan kereshetőek legyenek. Csak azokban a közgyűjteményekben lesznek elérhetőek tehát, amelyek erre megfelelő technikát alkalmaztak és kellő pénzt áldoztak.

B. S.: Azért a Google védelmében hozzátenném, hogy mostani szemmel is elképesztő dologra vállalkozott. Ha valaki 2005-ben azt mondta volna, hogy a világ összes könyvének digitalizálására vagy a világ összes útjának feltérképezésére készül, hihetetlennek hatott volna. Még a gondolat se jutott eszébe senki másnak. Ebből a szempontból meghajtom a fejemet a gondolkodók előtt. Másrészt, én egy picit jobban bírok abban, hogy a Google megőrzi az állományait, mint abban, hogy a Széchényi Könyvtár meg fogja őrizni azokat. Pusztul a digitális jelenünk, nem épül.

N. P. T.: Az valóban csak az egyik probléma, hogy a számítástechnika előtti világban keletkezett nyomtatott és nem nyomtatott kultúrvagyon hogyan őrződik meg. És tartok tőle, ez a kisebb kérdés. Látszik már, hogy az elmúlt hatezer évben főlhalmozott tartalmak belátható időn belül digitálissá válnak, és ettől kezdve az idők végezetéig az a könyvtárosok, levéltárosok és számítástechnikusok munkája lesz, hogy az ebben való keresgélést egyre jobba tegyék. Ez az egyik kérdés. De a másik kérdés, hogy a folyamatosan keletkező, naponta több terabájtnyi nagyságrendű digitális vagyon, milyen módon őrződik meg? És jelen pillanatban itt érzem a legnagyobb rést, hiszen ennek megőrzésére semmi nem kötelezi a kiadókat, de a kormányzati portálokat vagy internetszolgáltatókat sem.

L. L.: Én is nagy hiánynak érzem az interneten lévő digitális tartalmak elveszését, amely az internet működésétől kezdődően – Magyarországon 1992 óta – zajlik. Miközben a könyvtárak fő célja az elmúlt korszakok dokumentumainak megőrzése, ennek lényegét mintha nem értenék. Az első nyomtatott könyv, a Károli biblia hozzáférhető, de az első internetes weblap már nincs meg.

M. Zs.: Az elektronikus rendszerek fenntartása jóval drágább, mint a papíralapú rendszereké. Ahhoz informatikus és szerverterem kell, miközben a szerver-kapacitások folyamatos bővítést igényelnek.

N. P. T.: Az elektronikus adattárolás drágább lehet, mint folytonosan új és fenntartást igénylő épületeket építeni a papírhegyeknek? Kétlem. De még ha így is van, a papíron tárolt adat bármikor megsemmisülhet. Számos történelmi példánk van erre. Az elektronikus adatról viszont könnyű másolatot készíteni, amit akár egy másik kontinensen lehet tárolni...

M. Zs.: Az információk folyamatos pusztulásával együtt kell élnünk. Hosszú távon azonban a papírt tárolni olcsóbb, ez tény. Az elektronikus adattárolásnál nem elég, ha van egy szerver, arról tíz év múlva már nem tudunk adatot kinyerni. Az adatok folyamatos ellenőrzése és új és új formában történő elmentése is része a digitális megőrzésnek. Amíg a papírt fölteszem a polcra és nem csinállok vele semmit, csak pár percre villant kapcsolok, ha belépek a terembe, addig a szervernek meghatározott működési feltételek kellene, állandó ellenőrzés mellett. Ráadásul a hozzáférést is biztosítani kell. Az adatokkal dolgozó kutató számára szoftvert és számítógépet kell adnom az adatok megfelelő kezelésére. A papír esetében egy szék és asztal is elég.

E: Amennyiben technikailag lehetséges, meg kell tehát őrizni mindent?

L. L.: Eddig megőriztük.

B. S.: Én radikálisan azt mondom, pusztulnia kell a dolgoknak, mert nem tudunk megőrizni mindent. Az értéket őrizzük meg, ami felesleges azt engedjük el.

N. P. T.: Ne! Egy adott kor nem tudja megítélni azt, hogy 100 év múlva az érdeklődés, tudomány vagy a technológia mi mindennek a feldolgozását teszi lehetővé, s szükségessé. A 18. században a legtöbb nemzeti könyvtár nem tárolta el a napilapokat, a 19. században pedig a kinyomtatványokat, a levéltárak sok országban nem őrizték meg a népszámlálások egyedi számlálólapjait – senki nem látta előre, hogy a nagy sebességű számítógépek megjelenésével az egyedileg érdektelen adatok, szövegek korlátlanul újrakombinálhatókká, újraelemezhetővé válnak, s ezzel új társadalomtörténeti perspektívák nyílnak meg. A digitalizáció viszont azzal kecsget, hogy nem teljesen reménytelen átlátható terjedelemben megőrizni, ha úgy tetszik korunk teljes kultúráját. Hisz még ha ez több millió terabájt, akkor is elfér egy épületben.

A beszélgetést vezette és írásos anyagát szerkesztette Veroszta Zsuzsanna