



## Digitális gyűjtemények és adatok felügyelete: „kurátorság” és más alternatívák a tudományos könyvtárakban

A digitális anyagok sérülékenyek. Bár ez a nyomtatott dokumentumokra is elmondható, de azok helyes kezelésének módszereit kidolgoztuk az elmúlt évszázadok alatt, míg a digitális információ megőrzéséről sok mindent meg kell még tanulnunk. A „digitális kurátorság” egyre ismertebb fogalom: azoknak a tevékenységeknek a gyűjtőneve, amelyek a digitális adatok és tudományos eredmények karbantartására és használhatóságának megoldására irányulnak, azok teljes életciklusa alatt, a jelenben és a jövőben egyaránt. Magában foglalja a digitális archiválás és megőrzés folyamatát, de az adatok létrehozásának és menedzselésének helyes gyakorlatára is kiterjed, valamint azokra az értéknövelő eljárásokra is, melyekkel új információ és új tudás állítható elő. A hollandiai *Digitális Megőrzés Koalíció* koordinátora cikkében azt vizsgálja, hogy a tudományos könyvtárak milyen típusú digitális adatokat kezelnek, és hogy ezeket hogyan igyekeznek megőrizni:

- átadják-e ezt a terhet az erre szakosodott intézményeknek;
- vagy maguk vállalják fel a digitális megőrzés felelősségét, de a tényleges munkát kiszervezik és mással végeztetik el;
- vagy pedig teljesen magukra vállalják a feladatot és saját digitális repozitóriumot építenek ki az intézményen belül?

### Elektronikus publikációk: a raktározástól a hozzáférésig

A digitális kor egy sor új kihívást hozott a szakkönyvtárakba, ezek egyike az áttérés a nyomtatott könyvekről és folyóiratokról az elektronikus változatokra. Ez a váltás másfajta beszerzési gyakorlatot eredményez, például licenctárgyalásokat – jellemzően nagybani – előfizetésekről. És változik a hozzáférés módja is, mert többnyire az interneten át, online érhető el a tartalom. Ezáltal szinte észrevétlenül eltűnt a könyvtárak egyik alapfeladata: az információ raktározása. A könyvtárak szá-

mára meglehetősen nyugtalanító érzés, hogy nem birtokolják többé a tudományos publikációkat, hanem meg kell bízniuk a kiadóknak, illetve a közvetítőknél, hogy betartják a szerződés „folyamatos hozzáférés”-re vonatkozó pontját. A kiadók pedig egyre inkább a szaporodó digitális őrzőhelyekkel szerződnek, mint amilyen például a Holland Nemzeti Könyvtár e-Depot szolgáltatása, vagy a Portico és a LOCKSS; gyakran akár többel is, hogy az eltérő megőrzési technikákkal minél jobban csökkentsék az adatvesztés kockázatát. Más esetekben törvényi szabályozás alapján történik a digitális dokumentumok biztonságba helyezése – többnyire a nemzeti könyvtárakban.

Leginkább jól szerkesztett, PDF formátumban tárolt publikációkról van szó, melyek archiválása és a jövőbeli hozzáférés biztosítása nem különösebben nehéz műszaki feladat. Emellett az ilyen elektronikus kiadványok raktározása nem is elsősorban a tudományos könyvtárak felelőssége, hanem a kiadóké és a digitális megőrzéssel foglalkozó nemzeti vagy nemzetközi szervezeteké.

### Helyi digitalizált gyűjtemények

A megőrzendő anyagok másik csoportját a digitalizált gyűjteményrészek jelentik. Mivel a könyvtárak elkezdtek digitalizálni állományuk egyes darabjait, hogy így könnyebben hozzáférhetővé tegyék őket az olvasóknak, logikusnak látszik, hogy egyúttal azt a felelősséget is felvállalták, amit ezeknek a – sokszor jelentős költséggel létrehozott – számítógépes állományoknak a hosszú távú megőrzése jelent, hogy ezek a jövő generációi számára is használhatók maradjanak. Ám valójában ez nincs mindig így. A cikk szerzője számos olyan előadást hallott az éves LIBER-konferenciákon, amikor az előadó ezzel zárta a mondandóját: „Digitalizáltuk a gyűjteményt és ezáltal megőriztük a jövőnek. Köszönöm a figyelmet!” Az élet azonban nem ilyen

egyszerű. Nem lehet eleget hangsúlyozni, hogy a digitalizálás nem ugyanaz, mint a digitális megőrzés. Ha valamit számítógépes formára alakítunk, az csak az első lépés – keletkezik egy digitális objektum, ami ki van téve a jól ismert veszélyeknek:

- a hordozóeszköz tönkremenetele,
- a hardver vagy a szoftver elavulása,
- a szervezet megszűnése, és végül, de nem utolsósorban:
- emberi hibák.

Senki nem ismeri a végleges válaszokat ezekre a problémákra és valószínűleg nem lesz tökéletes megoldás később sem. Viszont a kurátorság már most is képes minimalizálni a kockázatokat egy olyan mozgó környezetben, ahol a digitális adatok folyamatosan özönlenek és a számítástechnika pedig állandóan változik. A gyakorlatban egy biztonsági mentés házon kívüli tárolása megfelelő kockázatcsökkentő megoldásnak látszik, de valójában csak egy óvintézkedés, ami nem nyújt védelmet az itt felsorolt veszélyek mindegyike ellen. Ha azt akarjuk, hogy a digitalizált anyag fennmaradjon, akkor folyamatosan gondozni kell. Itt nem létezik a „jótékony hanyagság”, mint a nyomtatott művek világában, amikor is örökre eltűntnek hitt antik könyvek csodálatos módon egyszer csak újra felbukkannak poros padlásokon. A digitális információ életben maradása teljesen a megfelelően működő hardver- és szoftverkörnyezet függvénye. Az *Ausztrál Nemzeti Archívum* ezért is vezette be a „teljesítőképesség-mutató” (performance model) fogalmát arra az esetre, amikor egy digitális objektumot elő kell hívni: objektum+(hardver+szoftver)=teljesítőképesség. Ennek a „teljesítőképességnek” a megvalósítása többet jelent annál, mint hogy biteket tárolunk egy szerveren. Olyan szervezetet igényel, amely nemcsak arra ügyel, hogy ezek a bitsorozatok ne sérüljenek, hanem folyamatosan figyeli a világban zajló technológiai változásokat, amelyek hatással lehetnek a megőrzendő anyagra és olyan stratégiákat dolgoz ki, amelyekkel megoldható, hogy ezek a digitális objektumok a következő generációs számítógépeken is megjeleníthetők lesznek. Ilyen jól ismert stratégiák például a migrálás, az emuláció/szimuláció és a normalizálás – ezek mind meglehetősen bonyolult műszaki eljárások. Kulcsrakész és mindenre kiterjedő termékek még nem léteznek a piacon ezekre a feladatokra. Vagyis: minden olyan könyvtárnak, amely saját maga akarja felügyelni és megőrizni a digitalizált gyűjteményét, bizonyos fokig késznek kell lennie arra, hogy részt vegyen az alkalmazott műszaki megoldások továbbfejlesztésében is,

különösen, ha egy már meglévő hozzáférési rendszert szeretne fenntartani.

Jó hír, hogy a saját digitalizált gyűjteményrészeknél rendszerint csak néhány fájlformátum fordul elő (PDF, JPG és TIFF), melyek kurátorsági szempontból viszonylag könnyen megőrizhetők hosszabb távon is. Egy másik fontos tényező, amely megkülönbözteti ezeket a digitális anyagokat a következő fejezetben tárgyalt tudományos kutatási adatoktól, hogy ezeket rendszerint maguk a könyvtárak hozzák létre, így feltehetően megfelelően strukturált és részletes metaadatok is tartoznak hozzájuk, ami megkönnyíti a velük való munkát és a megőrzést. De ezen könnyebbségek ellenére a hosszú távú megőrzés ebben az esetben is komoly vállalkozás egyrészt a költség miatt (ami nem annyira a tárolóeszközök árát jelenti, mert azok egyre olcsóbbak, hanem a szervezeti háttér fenntartását), másrészt a szükséges szakértelem miatt (mert bár már egyre több üzleti vállalkozás is megjelenik ezen a területen, az archívum üzemeltetése és összekapcsolása a könyvtárban működő információs rendszerekkel, sok helyi és számítástechnikai ismeretet igényel). Kérdés ezek után, hogy érdemes-e a tudományos könyvtáraknak felvállalni ezt a feladatot és felelősséget, különösen most, hogy a költségvetés egyre szűkebb és a helyzet a belátható jövőben sem lesz jobb? A nemzeti könyvtáraknak nem nagyon van más választási lehetőségük: többnyire törvény írja elő, hogy a nyomtatott és a digitális anyagok őrzőhelyeként kell működniük. De a szakkönyvtáraknak több alternatívájuk is van:

- Egyszerűen házon belül tárolják a digitális gyűjteményüket és remélik a legjobbakat. Talán ez a leggyakoribb megoldás, de kockázatos. Viszont ha az eredeti állomány épségben marad, akkor lehet azzal érvelni, hogy hosszú távon olcsóbb lesz újradigitalizálni egyes tönkrement fájlokat, mint gondosan őrizni a teljes digitális gyűjteményt. Ám ez csak akkor fogadható el, ha valóban a veszélyek alapos megfontolása után döntöttek így; vagyis például garantálni tudják, hogy a fizikai hordozókon levő eredeti dokumentumok sértetlenek maradnak a jövőben.
- Keresnek egy külső intézményt, akár egy nemzeti szervezetet, akár egy fizetős céget, és ott tárolják a digitális állományt. A tudományos könyvtárak egy részének ez járható út lehet, különösen amelyek túl kicsinek tartják magukat ahhoz, hogy saját digitális raktárt üzemeltessenek. Fontos viszont, hogy valóban megbízható őrzőhelyet találjanak, olyat, amely tényleg képes a szükséges felügyeletet ellátni, és amely gondoskodni tud ar-

ról, hogy a tárolt anyag integrálható legyen a könyvtári rendszerbe. Az adattárolásra szakosodott vállalkozások valójában gyakran csak biztonsági mentést képesek nyújtani, miközben hosszú távú őrzőhelyként hirdetik magukat. (Hasonló ez ahhoz, mint amikor úgy reklámoznak egy optikai tárolóeszközt, hogy ezer évig is megmaradnak rajta az adatok – ami jól hangzik, de ez semmi védelmet nem jelent a hardver és a szoftver elavulása ellen. Hiába maradnak rajta a bitek ezer évig, ha nem lesz olyan számítógép, amely fel tudná dolgozni őket.)

- A harmadik lehetőség egy saját digitális archívum kiépítése. Ez a leginkább nagyratörő és mélyreható választás a lehetséges alternatívák közül.

### **Digitális kutatási adatok**

Miközben a szakkönyvtárak megszabadultak a kiadói e-folyóiratok őrzésének felelősségétől, egy másik feladat jelent meg náluk: a kutatási adatok tárolása és gondozása. Ez igazi kihívás, tekintettel arra, hogy a tudományos adatok a legkomplikáltabb formáját jelentik a megőrzendő digitális információknak. Mind az adatok, mind az előállítók, rendkívül változatosak, és olyan komplex digitális objektumokat is tudni kell kezelni, mint az „élő” adatbázisok. Ugyanakkor a Google-szerű vállalkozások korában, amikor úgy tűnik, hogy a szakkönyvtárak lassan elveszítik az összes tradicionális életcéljukat, itt egy lehetőség arra, hogy új, mással nem pótolható helyet vívjanak ki maguknak – egyenesen az oktatási vagy kutatási anyaintézményük információs rendszerének közepében.

Jelenleg elég kevés egyetemi könyvtárnak van olyan digitális repozitóriuma, amely kellően feltöltött ahhoz, hogy valódi segítséget jelentsen a tudományos kutatásban és az oktatásban. Mindazonáltal a szakkönyvtáraknak legalább három olyan alapvető sajátosságuk van, amelyek különösen alkalmassá tehetik őket a tudományos kutatási eredmények kurátorságára:

- a küldetésükben benne van a hosszú távú megőrzés,
- strukturális támogatást kapnak,
- részei a tudományos közösségnek.

Igaz, hogy a második érv megkérdőjelezhető, mert a könyvtárak finanszírozása ritkán elégséges, viszont ők legalább költségvetésbe beépített támogatást kapnak. Mióta a tudományos kutatás projektrendszerben és határozott időre szóló pénzügyi

támogatásokkal folyik, a folyamatos finanszírozás hiánya miatt a kutatás során összegyűjtött adatok gyakran odavesznek, miután a projektek lezárulnak. Ennél még a könyvtárak szerény, de legalább biztos költségvetése is nagyobb garanciát jelent a hosszú távú megőrzéshez.

Ami azt az előnyt illeti, hogy a könyvtárak részei a tudományos közösség hálózatának: ez nem vehető biztosra. Ehhez a könyvtári és adatkezelési szolgáltatásokat speciálisan az adott kutatói kör igényeire kell(ene) szabni, és elsődleges prioritást adni az ő szükségleteiknek és kéréseiknek. Nem véletlen, hogy a digitális megőrzés területén leginkább használt Open Archival Information System keretrendszer kidolgozói kiemelt helyet és külön szakkifejezést rendeltek a „célközönséghez”.

Ezek a magas elvárások arra a döntésre is vezethetnek, hogy a tudományos könyvtár végül is nem száll be ebbe a folyamatba. Főleg, hogy a könyvtári szféra még inkább csak küzd a digitális világ különböző kihívásaival, miközben egyes tudományágak kutatói idejekorán felismerték a digitális jövőt és ügyesen megszervezték a maguk információs hálózatát, akár európai, sőt akár világméretben is. Ezek a nagy hálózatok általában jól megalapozottak és az adott tudományág területén dolgozók igényeire vannak optimalizálva. Így a tudományos könyvtáraknak nem sok szerep jut bennük, főleg mert a legnagyobb problémájuk a stabil és jelentős méretű pénzügyi háttér megszerzése, amiben a könyvtárak nem tudnak segíteni.

### **Tanácsok**

Befejezésül néhány gyakorlati, „tedd” és „ne tedd” típusú jó tanács azoknak a könyvtáraknak, amelyek azon gondolkodnak, hogy hogyan kezeljék a digitális gyűjteményüket és az általuk szolgált kutatók digitális információkra irányuló igényeit:

- Találjuk meg a „célközönségünket”.
- Üljünk le, írjunk egy tervet, és alakítsuk ki a célközönségünkkel kapcsolatos politikánkat. Sokszor már ennyi is elég, hogy felismerjük az erősségeinket és gyengeségeinket, és helyes sorrendbe tegyük a prioritásokat.
- A tervben vegyük számba azokat a digitális gyűjteményeket, amelyeket gondoznunk kell, és próbáljuk felbecsülni, hogy milyen problémákat okozhatnak a belátható jövőben.
- Beszéljünk a célközönségünk tagjaival az információs igényeikről és építsük be őket a tervünkbe.

- Selejtezzünk, és azután selejtezzünk újra. Bár néhány archiváló szakember még mindig az „őrizünk meg mindent” elvet hirdeti, mert „kik vagyunk mi, hogy eldöntsük, mire lesz szükség a jövőben?”, de a legtöbben már egyetértenek abban, hogy „mindent” megőrizni nem lehetséges és nem is célszerű.
- A méret számít. Az első digitális objektum, amit a kurátorságunk alá veszünk, rendkívül költséges lesz, de a milliomodik már csak fillérekbe kerül majd. Ha nem vagyunk elég nagyok ehhez, akkor válasszuk a következő pontot.
- Keressünk partnereket, lehetőleg hozzánk hasonló területen és mérettartományban. A gyakorlat azt mutatja, hogy az együttműködés leginkább a hasonló partnerek között jön létre, és egyre nehezkesebbé válik, minél távolabb vannak a szervezetek egymástól a hagyományaikat és a céljaikat tekintve.
- Válasszunk egy szakértői hálózattal rendelkező ernyőszervezetet. Ez lehet regionális, országos vagy nemzetközi.
- Kényeztessük el a célközönségünket. Engedjük, hogy a kutatók kutatók maradhassanak, például ne erőltessük rájuk a mi metaadat-sémáinkat, hanem inkább igyekezzünk nekik olyan adatkezelő eszközöket és megoldásokat kínálni, amiket könnyen be tudnak építeni a munkafolyamataikba.
- Döntsük el, hogy miben vagyunk jók, és ennek alapján keressük meg a helyünket az adatkurátori szervezetek hálózatában.

De ha ez nem megy, akkor legalább ennyit tegyünk meg:

- Soha ne tároljunk fontos digitális információt floppy lemezen, CD-n, vagy lokális PC-n, hanem töltsük fel megbízhatóbb hardverre.
- Rendszeresen mentjük az adatainkat és állapodjunk meg egy kollégánkkal, hogy vigyázzunk egymás back-up állományaira.
- Készítsünk listát a felügyeletünk alá tartozó digitális anyagokról és tartsuk naprakészen.
- Keressünk megbízható őrzőhelyet az adatainknak. Ez lehet a nemzeti könyvtár, vagy egy nemzeti archiváló szervezet, vagy akár egy olyan kolléga, aki digitális repozitóriumot üzemeltet. Ha nem áll módunkban hivatalosan meggyőződni arról, hogy egy archiváló hely mennyire megbízható, nézzük meg az intézmény egészét és tegyük fel magunknak a kérdést: valószínű, hogy ez a szervezet még ötven vagy száz év múlva is létezni fog?
- Az erőforrásainkat a hozzáférés kérdésére összpontosítsuk, mert hosszú távon csak az fog számítani. Minden más csak eszköz ahhoz, hogy a hozzáférés lehetőségét megőrizzük.

**/ANGEVAARE, Inge: Taking care of digital collections and data: 'Curation' and organisational choices for research libraries. = LIBER Quarterly, 19. köt. 1. sz. 2009. p. 1–12./**

*(Drótos László)*