



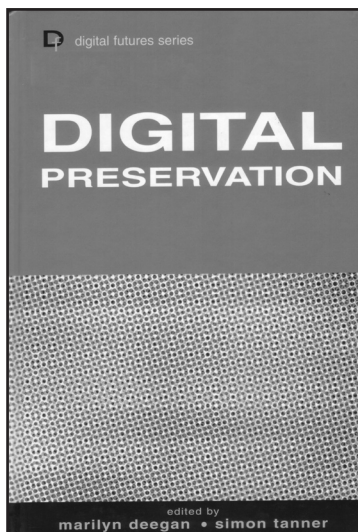
A digitális megőrzés kérdései

A szerkesztők (Deegan és Tanner) neve nem ismeretlen a Könyvtári Figyelő olvasói előtt, hiszen a *Digital Futures* elnevezésű sorozat első kötetéről már jelent meg ismertetés a lap hasábjain.¹ A digitális megőrzés kérdéseit tárgyaló kötet a sorozat harmadik tagja.

Bevezetőjüket azzal kezdik, hogy meglepőnek tűnhet, ha manapság valaki nyomtatott könyvet jelentet meg erről témáról. Hozzáfüzik azonban, hogy az összefüggő és átgondolt érvelés megfelelő fóruma ma is a nyomtatott könyv, amely hosszú távú megőrzésre is alkalmas, időálló formátum. Marilyn Deegan és Simon Tanner jegyzik a kulcskérdéseket áttekintő első fejezetet is. Felvetik, hogy sokak szerint a digitális megőrzés világosan megkülönböztethető a megőrzés más, hagyományos formáitól. Ez azonban csak részben van így, ezért a megőrzési kötelezettség fogalmait és alapelveit mindkét területen érdemes figyelembe vennünk.

A digitális megőrzés úgy definiálható, mint az a folyamat, amelynek során olyan környezetet hozunk létre és tartunk fent, amely a művek tárolása és/vagy használata szempontjából a legjobb, és az szolgálja, hogy megakadályozzuk e művek sérülését vagy romlását, továbbá lehetővé tegyük, hogy a lehetséges leghosszabb élettartamot éri meg.

A digitális megőrzés izgalmas kérdése a szurrogátumok használata vagyis az, amikor az eredetit egy olyan objektummal helyettesítjük, amelyik szimulálja annak tartalmát. A szurrogátumok hitelessége és elfogadásuk a felhasználók részéről legalább akkora probléma, mint maguknak a szurrogátumoknak a megóvása és konzerválása. Ráadásul a megőrzés sokkal több



Digital preservation / [ed. by] Marilyn Deegan and Simon Tanner.– London : Facet, cop. 2006.– XXIII. 260 p. – (Digital futures series)
ISBN 978-1-85604-485-1

figyelmet kapott, mint a hitelesség biztosítása. Szurrogátumokkal nem tudunk mindent pótolni, viszont szurrogátumok nélkül mindent elveszthetünk „romlékony” dokumentumok esetében.

Felmerül annak a kérdése is, hogy a digitális szurrogátumok helyettesíthetik-e a kipróbált analóg módszereket. A mikrofilm élettartama várhatóan 500 év, és van már ismeretünk romlásának mértékéről és tárolásának költségeiről. Nagyon kevés könyvtáros selejtezné az eredetit és bízna a digitális másolatban anélkül, hogy biztonsági mikrofilm-másolatot készítené. Minden digitalizálási projekt során érdemes megfontolnunk tehát, készítsünk-e mikrofilmet is. Vannak persze az eredeti dokumentumoknak olyan vonásai, amelyeket a mikrofilm nem tud megőrizni, digitalizálással viszont rögzíthetők.

A dokumentumok mellett az adatok megőrzése is fontos, mivel azokat mulékony hordozóra rögzítették, meghatározott fájlformátumban és átmeneti kódolási séma (programnyelv) szükséges ahhoz, hogy interpretálni tudjuk őket. A másik probléma az adatok összetettsége, valamint az, hogy az adatokból nyerhető jelentés attól is függ, hogy miként kapcsolták össze az egyes adatokat. Tudjuk, hogy a szöveges dokumentumok is igen összetettek, szerkezetüket azonban nem kell ahhoz megérteni, hogy meg tudjuk őrizni őket, csak interpretációjukra van szükség. Előfordulhat, hogy idővel elfelejtjük, hogyan lehet őket interpretálni, ennek fellevenítéséhez azonban csak emberi képességek szükségesek. A helyzetet bonyolítja, hogy a digitális adatok esetében az adatok és az emberi értelmezés közé be kell iktatni a számítógépet.

Az adatok megőrzésének két olyan kulcskérdése van, amelyeknek meglepően kevés köze van az eredeti bitek és bájtok (bitfolyamok) megőrzéséhez:

- a fizikai hordozó megőrzése,
- az interpretáció, az olvasás és a felhasználás eszközeinek megőrzése.

A digitális adatok megőrzése azért sürgető, mert ezek természetes élettartama nem elég hosszú ahhoz, hogy várjunk új hordozók megjelenésére, ráadásul az adatok önmegőrző képessége rossz.

A digitális megőrzésnek közismerten számos módszere van. A technológia megőrzése, vagyis az a módszer, amikor a digitális forrást támogató hardver- és szoftver-platformokat őrizzük meg, meglehetősen nehezen kivitelezhető és költséges. A frissítés technikai szempontból egyszerű, és kicsi a kockázata annak, hogy adatokat veszítünk el, ráadásul más megőrzési módszerekkel együtt alkalmazható. A migrálás és a technológiák emulációja pedig viszonylag széles körben ismert és használt megőrzési formák. Ezek mellett számításba jöhet az adatrégészet is. Ez akkor válhat szükségessé, ha olyan digitális forrást kell megmentenünk, amely fontos információt tartalmaz és valamilyen váratlan esemény következett be.

Megőrzési stratégiaként a hardver emulációját helyettesíthetné. Nem őriznénk meg a programokat és nem is emulálnánk őket. Inkább az adatregészek fejtenék meg az adatstruktúrákat, ami olcsó, viszont kockázatos megoldás.²

A digitális megőrzés stratégiai kérdés, tehát különösen fontos stratégiáinak áttekintése. Ezzel kapcsolatban mindig felmerül annak kérdése, hogy mi az az információ, amelyet valamikor használni fognak, és melyik az, amelyiket sohasem. Mivel jósolni nem tudunk, a legjobb, ha a lehető legalacsonyabb szintre igyekszünk levinni a megőrzés költségeit, és nem nagyon figyelünk oda arra, hogy vannak olyan anyagok, amelyeket nem fognak a jövőben használni. Egy könyvelő mondhatja, hogy tudomása szerint egy adott archívum tartalmának 90%-át sohasem fogják használni, annak törlésével tehát megtakarítás érhető el. A digitális gyűjtemény kurátora erre azt válaszolhatja, hogy csak azt mondják meg, melyik az a 90%, mert azt azonnal törölni fogja.

A stratégia szempontjából az elvonatkoztatás életfontosságú. A megőrzés problémája ilyen nézetben két feladat megoldását jelenti:

- Biztosítsuk, hogy rendelkezésre álljon az a hordozó, amelyik lehetővé teszi a régi technológiától az újabb technológiába irányuló bitfolyam tárolását és másolását.
- Gondoskodjunk róla, hogy fennmaradjon az adatformátumok ismerete és képesek legyünk arra, hogy ezeket az adatformákat gazdaságosan dolgozzuk fel.³

Aligha véletlen, hogy a kötetben viszonylag előre került a metaadatok kérdése, különösen azért, mert a metaadatok sokfélék lehetnek. A – bitek sorozataként tárolt – digitális objektumok integritásának megőrzése jól ismert folyamat, amely kevés metaadatot igényel. Az, hogy hozzá is férhessünk a digitális tartalomhoz, már jóval összetettebb feladat. Szükséges az is, hogy ez a tartalom értelemmel bírjon. Tudnunk kell továbbá, hogy módosították-e a dokumentumot, és ha igen, ki, mikor tette ezt. Ez utóbbi dokumentálást olyan metaadatok szolgálják, amelyeket többféleképpen, például származási metaadatoknak

neveznek, és természetesen leíró metaadatokra is szükség van.⁴

Két írás is foglalkozik a web archiválásával, az elvi kérdések kifejtése mellett esettanulmányokat⁵ is felvonultatva. *Julien Masanès* írásában kifejti, hogy a weboldalak archiválása nem magától értetődő feladat. Először is az adott szájt összes oldalát meg kell találnunk, bár a http-protokoll nem ad teljes listát ezekről. A begyűjtésre használt automatikus eszközök (crawlerok) számára egyes linkek feloldása nehézséget jelenthet. A láthatatlan (deep) web megléte is okozhat problémákat. Az időtényező sem lényegtelen, hiszen például egyes belinkelt oldalak tartalma megváltozhat, és eltérhet attól, mint amilyen akkor volt, amikor a rá mutató oldalt begyűjtöttük. A webarchívumok kialakításakor ráadásul problémát jelent, hogy egy-egy weboldal nem olyan, mint egy (nyomatott vagy digitalizált) könyv, amely véges számú oldalból áll, és amelynek meghatározott sorrendje⁶ van. A költségek kérdése két fejezetet kapott. *Brian F. Lavoie*, aki – a szerkesztőket nem számítva – a neves szerzőgárda talán legismertebb tagja, többek között a bizonytalansági tényezőkről ír. Ezek egyike az, hogy nincs megegyezés arról, hogy melyek a megőrzés legjobb formái, milyen a bevált gyakorlata.

A másik bizonytalansági tényező a megőrzési idő. A kulturális örökség intézményei elkötelezettsége ebben a tekintetben nyitott végű, mivel örök időnkig kívánják megőrizni a rájuk bízott digitális objektumokat. Nehéz viszont megbecsülni, hogy például a jelenleg használt hardverek és szoftverek mikor fognak elavulni. Pedig a digitális megőrzés létjogosultságát éppen az adja meg, hogy a technológiai környezet folyamatos változása az eszközök elavulása folytán veszélyezteti a digitális objektumok fennmaradását. Maguk a megőrzés céljai is bizonytalanok, azaz közel sem olyan egyértelműek, mint az analóg dokumentumok esetén. Az előbbieket megőrzésének formája különbözhet eredeti formájuktól. A megőrzés, formájától függetlenül jelentős és folyamatosan felmerülő költségeket jelent. Nemcsak előzetes beruházásokra van tehát szükség,

hanem a repozitóriumok fenntartásának anyagi feltételeit is biztosítani kell.

A digitális megőrzés egyre inkább folyamatos tevékenységgé kell, hogy váljon. Proaktív módon meg kell előznie a károsodásokat és az elavulást, tehát nem csak reagálnia kell a már bekövetkezett elavulásra.⁷

Stephen Chapman írásának beszélő címet adott: „A hosszú távú megőrzésben a pénz számít”.⁸

A kötet záró fejezetei közül az első a digitális megőrzéssel kapcsolatos európai uniós megközelítéseket mutatja be⁹. Ezt követi az a fejezet, amely számos projektet elemez esettanulmány formájában¹⁰.

A könyv, bár többszerzős, nem tanulmánygyűjtemény, hanem egységes mű kíván lenni, amit alátámaszt az is, hogy egyetlen, egységes irodalomjegyzék járul a fejezetekhez.

Érdeemes viszont a könyvvel kapcsolatos egyik külföldi recenzió néhány megállapításával ki-

egészítenünk az eddig mondottakat, melyekkel messzemenően egyetérthetünk, ráadásul ez az ismertetés a tekintélyes *Journal of Documentation* hasábjain jelent meg. Azt olvassuk benne, hogy a könyv hasznos bevezetés a digitális megőrzés világába. A fejezetek hossza és részletessége változó, emiatt a kifejtés néha nem eléggé informatív és nehezen olvasható, miközben jól ismert szakértők írásait olvashatjuk, ami önmagában is érdem és egyúttal meg is magyarázza a fejezetek egyenetlenségét és azt, hogy a szerkesztők által írt bevezető tanulmányon kívül az írások nem alkotnak koherens egészet, ilyen módon a mű a klasszikus értelemben nem is nevezhető kézikönyvnek. Ez a könyv nem a szakértőknek szól, hanem azoknak, akik egy keveset már tudnak a megőrzésről, viszont tudásukat bővíteni szeretnék.¹¹

Koltay Tibor

Jegyzetek

- 1 Deegan, M., Tanner, S.: Digital futures. Strategies for the information age. London: Library Association Publishing, 2002. Ismerteti Koltay Tibor: A digitális jövő. = Könyvtári Figyelő, 49. évf., 2003. 4. sz., 837–841. p.
- 2 DEEGEAN, M. – TANNER, S.: Key issues in digital preservation. In: Digital preservation. 1–31. p.
- 3 HOLDSWORTH, D.: Strategies for digital preservation. Uo. 32–59. p.
- 4 WENDLER, R.: The status of preservation metadata in the digital library community. Uo. 60–77. p.
- 5 MASON, E.: Web archiving activities: case studies. Uo. 98–105. p.
- 6 MASANÉS, J.: Web archiving. Uo. 78–97. p.
- 7 LAVOIE, B.F.: The costs of digital preservation. Uo. 106–132. p.
- 8 CHAPMAN, S.: It's money that matters in long-term preservation. Uo. 133–146. p.
- 9 MCKINNEY, P.: Some European approaches to digital preservation. Uo. 147–165.
- 10 KELLY, J. – MASON, E.: Digital preservation projects: some brief case studies. Uo. 166–217. p.
- 11 VILAR, P. c.n. = Journal of Documentation, Vol. 64. 2008. No. 4. 628–630. p.

A JISC (Joint Information Systems Committee), a brit felsőoktatási informatikát finanszírozó testület bejelentette, hogy 2010 augusztusától nem tudja finanszírozni teljes egészében az Intute (<http://www.intute.ac.uk>) multidiszciplináris forráskalauz működését. A JISC mindig törekedett arra, hogy az általa támogatott intézmények alternatív finanszírozási forrásokat is keressenek. A jelenlegi gazdasági helyzetben azonban erre nincs reális lehetőség. Az Intute augusztus elseje után is elérhető lesz, de karbantartását minimális mértékben fogják elvégezni.

(Eredeti forrás: <http://www.intute.ac.uk/blog/category/intute-news/>;

Katalist, 2010. febr. 28. Koltay Tibor híre)