

A digitalizálás ellenőrzése: az eddigi tapasztalatok tanulságai

Bevezetés

A digitalizálás a kulturális emlékezetet őrző intézmények, vagyis a könyvtárak, levéltárak és múzeumok gyűjteménygondozó tevékenységének fontos eleme, mivel széles körben elérhetővé teszi a gyűjtemények darabjait kutatási, oktatási, művelődési, idegenforgalmi vagy egyéb célokra, és emellett sok esetben az eredeti példány védelmét is szolgálja, hiszen így kevesebbszer kell azt kézbe venni. Ma már olyan széles körű nemzetközi együttműködések formájában zajlik ez a tevékenység, mint például a *European Library*, melyben 48 nemzeti könyvtár vesz részt, illetve az *Europeana*, amely az EU országainak közgyűjteményeiben digitalizált kulturális örökség nyilvántartása. Az *Európai Bizottság* az FP7, vagyis a Hetedik Keretprogram projektjei esetében 2007–2008 között 40 millió eurót, az eContentplus keretében pedig a 2005-től 2008-ig terjedő időszakban 60 millió eurót szán a digitalizálás és a digitális könyvtárak támogatására. Emellett természetesen nemzeti szinten is komoly forrásokat fordítanak a digitalizálásra: egy közelmúltban közzétett EB- jelentés szerint például Görögországban 100 millió eurót költöttek ilyen projektekre 2003–2007 között, míg Hollandiában csak az audiovizuális gyűjtemények digitalizálására 90 millió eurót terveztek be a 2007–2014 közötti időszakra. Az is gyakori, hogy az üzleti szféra is bekapcsolódik a nagy tömegű digitalizálási munkákba és azok finanszírozásába – elég, ha csak a Google 2004-ben indult könyvszkennelő projektjét, vagy a *NARA (US National Archives and Records Administration)* és az *Amazon* filmdigitalizáló vállalkozását, vagy az *Open Content Alliance* együttműködést (Microsoft, Adobe, Yahoo!) említjük példaként.

Nem meglepő tehát, hogy mindinkább előtérbe kerül ezeknek a komoly összegeket felemésztő és nagy jelentőségű digitalizálási munkálatoknak a szisztematikus és folyamatos felügyelete, hogy ellenőrizni lehessen a minőséget, a hatékonyságot és a pénzek gazdaságos felhasználását. A straté-

giák alakítói, a finanszírozók és az intézményvezetők számára egyaránt fontos, hogy rendszeresen begyűjtött és kiértékelt, egymással összevethető statisztikai adatokat kapjanak erről a területről. Annak ellenére, hogy növekszik az igény a digitalizálási munkák monitorozása iránt, még csak kezdeti próbálkozások történtek eddig: nincsenek nemzetközileg elfogadott szabványok és ajánlások a szükséges információk körére és megszerzésük módjára vonatkozóan. Jelen tanulmányában a litván szerző az eddigi fontosabb felmérések elemzése alapján összegzi a levonható tanulságokat, a helyes és a hibás módszertani megoldásokat, segíteni próbálva ezzel is a megfelelő nemzeti és nemzetközi modellek kialakítását a digitalizálási munkák előrehaladását jelző statisztikai adatok begyűjtése terén. Az elemzés kiterjed a felmérések módszertanára, a digitalizálási célokra, továbbá a digitális gyűjtemények használatának, méretének és növekedési ütemének mérésére, valamint a finanszírozási módokra és a költségekre vonatkozó kérdésekre.

A felügyeleti rendszer kereteit alapvetően meghatározza az, hogy hogyan definiáljuk magát a digitalizálást. Ha csak technikai értelemben nézzük, akkor a digitalizálás fogalma az analóg tartalom digitálissá, vagyis számítógéppel feldolgozható bináris kóddá való alakítását jelenti. Ez a fajta, információmenedzsment-szemponitú megközelítés egységes folyamatokat jelent a kulturális örökséget alkotó mindenféle gyűjtemény esetében – legyen az könyvtár, levéltár vagy múzeum: vagyis a digitalizálendő anyagok kiválasztását és előkészítését, majd ezek számítógépes formára való alakítását és gyűjteménybe integrálását, valamint tárolását és a hozzáférés megoldását. A digitalizálás hatása azonban szélesebb körű és hosszabb távú, mint maga a digitalizálási folyamat, ezért stratégiai menedzsmentkérdéseket is felvet: például projektervezés, kockázatkezelés, kiszervezés, partnerkapcsolatok építése, finanszírozási problémák, minőségbiztosítás, emberi erőforrások menedzselése. Van egy harmadik fajta, szolgáltatásközpontú

megközelítés is, amely a digitalizálást csak egy eszköznek tekinti ahhoz, hogy az adott intézmény minél jobban elláthassa a küldetését és kiszolgálhassa a felhasználóit. Mivel céljaikat és felhasználói elvárásait tekintve a különféle intézmények jelentősen eltérhetnek egymástól, ezért ha az általánosan érvényes állományvédelmi és hozzáférésbiztosítási indokok helyett a szolgáltatásokat helyezzük előtérbe, akkor a monitorozási szempontok is eltérőek lesznek intézménytípusokként.

Előzmények

Az európai országokban folyó digitalizálási munkák adatainak összegyűjtése már a 2001-ben közreadott „Lundi alapelvek”-ben és „Cselekvési terv”-ben is stratégiai fontosságú feladatként szerepelt, de még nem született meg az ehhez szükséges átfogó módszertani keret. A témában megjelent publikációk is inkább csak egyes részkérdésekkel foglalkoztak eddig. Leggyakrabban a digitalizálási költségeket elemezték és különféle költségkalkulációs modelleket állítottak fel. Egy másik népszerű kutatási terület a digitális tartalomszolgáltatások használatának mérése: például virtuális látogatások, keresések és letöltések száma. Az ezekkel kapcsolatos indikátorok és mérési módszerek az ISO 11620 és az ISO 2789 szabványokba is bekerültek, illetve az európai EQUINOX valamint az amerikai LibQUAL könyvtári minőségügyi projektekben is megjelentek. A digitalizáláshoz kapcsolódó statisztikai adatok és teljesítménymutatók egységesítése terén eddig a könyvtárak érték el a legnagyobb előrehaladást; a múzeumok most teszik meg az első lépéseket ezen az úton, míg a levéltárak esetében nem látszanak még ilyen standardizálási törekvések. Olyan kezdeményezés pedig egyáltalán nem volt még, amely nemzetközi méretekben is egységes statisztikai keretrendszerrel dolgozott volna ki valamennyi gyűjteménytípusra.

Kutatási módszer

A cikkben ismertetett elemzés céljára 32 többnyire országos vagy nemzetközi szintű statisztikai felmérés kérdőívét és összefoglaló jelentését választották ki, amelyek nagyrészt olyan szakmai, illetve kormányzati szervezetek honlapjairól származtak, mint az IFLA, az *International Council of Archives*, az *International Council of Museums*, az UNESCO, valamint a kulturális örökség ügyében illetékes minisztériumok. A felmérések egy része kifejezetten a digitalizálási munkákra terjedt ki (15 jelen-

tés), mások ennél átfogóbb témájúak voltak (13), és akadt néhány, amely csak a digitalizálás valamely szűkebb részterületére vonatkozott. A jelentések szövegét kvantitatív és kvalitatív tartalomelemzésnek vetették alá, vagyis a könnyen számszerűsíthető jellemzőkből kimutatható trendek mellett induktív módszerrel – tehát nem előre definiált kategóriákat használva – is igyekeztek mintákat találni a szövegekben. Az így kialakított szemantikus kategóriák eloszlási gyakoriságát azután szintén elemezték és ábrázolták.

Eredmények

Elsőként az *általános módszertani elveket* vették górcső alá, például azt, hogy a felméréseket végzők hogyan választották ki a megkérdezettek körét, és hogy milyen digitalizálási részterületekre vonatkoztak a kérdéseik. Földrajzi szempontból a felmérések 34%-a volt nemzetközi, 44%-a országos jellegű és 22%-a kisebb, csak néhány intézményre kiterjedő. A legtöbb kutatásnál olyan intézményeket kerestek meg, amelyekről feltételezték, hogy folyik náluk digitalizálás – ezek többnyire könyvtárak, levéltárak és múzeumok voltak, de előfordultak köztük kormányzati szervek és az üzleti szféra képviselői is. Más esetekben a szakmai szervezetekben vagy konzorciumokban való tagság, illetve a közös tevékenység vagy gyűjtőkör volt a kiválasztás alapja. Csak néha előzte meg szisztematikus válogatás a címzettek listájának összeállítását.

Ami a másik kérdést, a digitalizálási folyamat egyes fázisait illeti: mindössze hat jelentésben volt definíció a digitalizálás fogalmára, és bár a kérdőívek általában többféle részletkérdésre is kiterjedtek, gyakran nem volt semmilyen indok a jelentésekben, hogy miért pont ezeket a területeket választották ki a kutatók, miért ezeket tekintették fontosnak a digitalizálás sikerességének megítélése szempontjából. Két nagy csoportba sorolhatók azok a részterületek, amelyek a felmérésekben szerepeltek: stratégiai menedzsment és információforrás-menedzsment. Előbbihez olyan kérdések tartoznak, mint a digitalizálási politika kialakítása, a fenntarthatóság, a pénzügyi, emberi és technikai erőforrások megteremtése; utóbbihoz pedig a digitális tartalom előállítás, feldolgozása, tárolása, a hozzáférés biztosítása stb. Maguk a kérdőívek is kétféleképpen bizonyultak, attól függően, hogy milyen szintű szervezet kezdeményezte a kutatást. A felmérések 31%-a minisztériumok vagy más kormányzati szervek megrendelésére készült és ezek

tipikusan *top-down*, vagyis felülről lefelé irányuló megközelítést alkalmaztak. Elsősorban az olyan stratégiai témákra helyezik a hangsúlyt, mint a projektek száma, az együttműködések, a finanszírozás, a célközönség, az esetleges akadályok; az információforrás-menedzsment jellegű kérdések közül a digitalizált dokumentumok típusa és mennyisége, a hozzáférés módja és a használat mértéke iránt érdeklődtek leginkább. A 69%-ot kitevő *bottom-up* típus a szakmai szervezetek, a kutatóintézetek, a konzorciumok által indított felmérésekre volt jellemző. Ezek olyan szempontokat helyeztek előtérbe, mint például a digitalizálási költségek, a munkák kiszervezése, a dokumentumok kiválasztása, a digitális formátumra való konvertálás, a tárolás, a copyright, a keresőeszközök, a metaadatok. A kérdőíveken megjelenő témák elemzése azt mutatja, hogy a kérdéseket meglehetősen szubjektív módon, a korábbi tapasztalatokra alapozva és a felmérést végzők szerepének függvényében állították össze, nem pedig annak alapján, hogy milyen tényezők fontosak valójában egy digitalizálási projekt sikerességének megítélésében.

18 kutatásban kérdeztek rá a *digitalizálás céljára vagy indítékára*, de ezt többször összekeverték a kiválasztási szempontokra vonatkozó kérdésekkel, pedig az, hogy miért digitalizálunk, egy általánosabb dolog, mint az, hogy miért pont azt a dokumentumot vagy tárgyat választottuk ki digitalizálásra – bár az előbbi természetesen befolyásolja az utóbbit. A digitalizálási célra adható válaszokat három jellemző csoportba lehetett sorolni: a hozzáférés elősegítése (pl. oktatási vagy kutatási célokra), belső gyűjteménygazdálkodás (pl. dokumentumszolgáltatás, raktári tárhely megtakarítása, bevételszerzés), és megőrzés (pl. állományvédelem, elavult hordozókról való migrálás). A felhasználók igényei alulreprezentáltak voltak a célok és a motivációk között, inkább az intézmény saját érdekei kaptak prioritást, ami ellentmondásban van a digitalizálást ösztönzők és a finanszírozók elvárásaival, mint például a nemzeti identitás megőrzése a globalizált világban, a kultúrák dialógusának elősegítése, a tudomány és az oktatás támogatása, és egyben leszűkíti a digitalizálási munkák eredményeinek hatását és hatókörét.

A felhasználói igény, a használhatóság és hozzáférhetőség 11 kutatásban szerepelt a kérdések között, jellemzően ilyen módon: virtuális látogatók és látogatások száma, illetve felhasználói elvárások és ezek mérési módszerei. Egy kifejezetten erre a témára irányuló 2003. évi IMLS kutatás eredményei azt mutatták, hogy a végfelhasználók

igényeinek meghatározása gyakran az ő közvetlen bevonásuk nélkül történik. Sokszor inkább helyi vagy külső szakértők, esetleg egy tanácsadó bizottság döntött a digitalizálási projekt tartalmáról és megvalósítási módjáról. Kevés helyen alkalmaztak fókuszcsoporthoz vagy közvéleménykutatásos módszereket, ott is inkább csak a végeredmény tesztelésére, nem pedig a fejlesztési folyamat minden fontos fázisában. Erre az egyik magyarázat az lehet, hogy még kevesen ismerik a használhatóság mérésének módszertanát, sőt egyes válaszokból az is kiderült, hogy néhol teljesen félreértették a „*usability*” (használhatóság) fogalmát. Az is igaz, hogy míg a használat mértékének, vagyis a forgalomnak a méréséről már sok szakmai vita lezajlott, a használhatóság és a hozzáférhetőség megállapításához nem állnak még rendelkezésre általánosan alkalmazható ajánlások a kulturális örökséget őrző intézmények számára.

A projektfinanszírozás és a digitalizálási költségek volt a következő elemzés alá vont kérdéscsoport. A felmérések közül 11 foglalkozott a finanszírozás témájával: elsősorban a pénzforrásokra kérdeztek rá, de néhány esetben ezek elérhetőségét, elégségességét és fenntarthatóságát is firtatták. A források esetében a kérdőíveken felkínált válaszok hét csoportba voltak sorolhatók: leggyakrabban a belső finanszírozás, vagyis a közgyűjtemény vagy anyaintézménye saját költségvetéséből fizetett digitalizálás fordult elő, emellett pedig az állami és alapítványi pénzek felhasználása, továbbá a szponzori támogatásból, illetve a fizetős szolgáltatásból való fenntartás szerepelt még többször. Néhány esetben arra is megkérték a válaszadókat, hogy becsüljék meg az éves digitalizálási költségeiket, sőt olyan felmérés is akadt, ahol ezeket részterületekre bontva kellett megadni: például szakértői díjak, a digitalizálók fizetése, szerzői jogdíjak, marketing- és reklámköltségek. Egy 2006-os kutatásban, amely a felsőoktatási könyvtárak és a szakkönyvtárak digitalizációs tevékenységét igyekezett felmérni, megkülönböztették az indulási és a fenntartási költségeket is, hiszen a szükséges műszaki infrastruktúra kiépítése jelentős beruházás, míg a folyamatos munka már nem annyira költséges – és természetesen a későbbi alacsonyabb ráfordítás nem jelenti azt, hogy az intézmény ilyenkor már kevésbé elkötelezett a digitalizálás iránt. A költségvetések részletes megismerése hasznos információ, és következtetéseket lehet levonni belőle a menedzselés hatékonyságára is, de a tapasztalatok szerint nem könnyű pontos és egymással is összevethető adatokat szerezni az intézményektől. Részben mert nem biztos,

hogy a válaszadók hozzáférnek a szükséges pénzügyi adatokhoz, részben pedig mert néha nehezen választhatók külön a digitalizálás költségei más tevékenységeiktől (pl. reprográfia, illetve kiállítások vagy kiadványok anyagának előkészítése), vagy mert nincs is erre külön költségvetési keret. A fajlagos digitalizálási költségek megállapítására csak öt kutatásnál tettek kísérletet, de ezekben sem határozták meg pontosan, hogy hogyan kellene kiszámítani az egy könyvre, egy folyóiratszámra, vagy egy oldalra eső költséget, úgyhogy az így nyert adatok szintén nehezen hasonlíthatók össze az egyes intézmények vagy projektek között.

A *digitális gyűjtemények méretére és gyarapodási ütemére* vonatkozó kérdéseket kvantitatív elemzésnek vetették alá. Eszerint a digitalizált anyagok fajtájára 15 kutatásban kérdeztek rá, és összesen 223 féle kategóriát adtak meg a kérdőívek összeállítói. Az osztályozási szempontok is változatosak voltak: például az információ megjelenési formája (szöveg, kép, hang), a publikáció típusa (újság, folyóirat, könyv), a dokumentum létrehozója (kormányzati kiadványok, személyes iratok), a mű témája (művészet, vallás, tudomány); sőt az *European Digital Library Project* esetében még földrajzi és időszak szerinti kategóriákat is megadtak. A digitalizálni tervezett és már digitalizált tételek mennyiségét vagy ezek arányát tíz esetben igyekeztek megtudni a kutatók, és 11 jelentésben szerepelt a digitális gyűjtemény aktuális méretére vonatkozó adat. Ez utóbbi meghatározására eltérő mértékegységeket használtak a kérdőíveken: leggyakrabban a „tétel” vagy „objektum” szót (de hogy mit kell egy tételnek/objektumnak tekinteni, azt rendszerint nem definiálták), máskor pedig a digitális képek, oldalak, vagy fájlok számát kellett megadni, vagy a gyűjtemény tárhelyfoglalását gigabájtokban, sőt olyan is előfordult, hogy a válaszadó maga választhatta meg a számára megfelelő mértékegységet. A leghasznosabb megoldás az, ha az analóg dokumentumokkal azonos egységben határozzák meg digitalizált tételek mennyiségét, mert ebből következtetni lehet a projekt méretére és előrehaladási ütemére, valamint a várható hasznosságára, míg a képek, fájlok vagy gigabájtok száma nem sokat jelent, és nem összevethető mértékegység, mivel ezek az adatok nagyban függenek a dokumentumok típusától és a digitalizálás módjától.

Tanulságok, következtetések

A digitalizálási felmérések elemzése azt mutatta, hogy ezek szinte mind felderítő jellegű, egyedi próbálkozások voltak, amelyeket nem a rendszeres adatgyűjtés szándékával terveztek meg, és nincs mögöttük kidolgozott módszertan. Sokszor még az sem volt egyértelmű, hogy a felmérés végző mit ért a digitalizálás fogalmán, a megkérdezett intézmények körét pedig feltételezések alapján, leginkább a típusuk szerint választották ki, nem pedig a kulturális örökség digitalizálásában valóban fontos és illetékes szereplőket keresték meg. A kérdőíves vizsgálatot kezdeményező szerv jellege nagyban meghatározta, hogy inkább a stratégiai, vagy inkább az információmenedzselési kérdések kaptak nagyobb hangsúlyt. Pedig mindkét témacsoport egyaránt fontos annak megítéléséhez, hogy egy intézmény a digitalizálás terén mennyire elkötelezett és kompetens, mennyire haladt már előre a digitalizálási célok megvalósításában és hogy hosszabb távon is képes lesz-e az ez irányú feladatait ellátni. Ahhoz persze, hogy valóban objektív és minden lényeges részterületre kiterjedő kérdőívet lehessen összeállítani, szükség lenne egy elfogadott modellre, amely pontosan meghatározza, hogy mik tartoznak a digitalizálási munkálatok közé, ezeknek mi a célja és elvárt eredménye, és mikor tekinthetünk sikeresnek egy digitalizálási projektet. A kérdőívek elemzéséből az rajzolódott ki, hogy a felhasználók jobb kiszolgálása helyett inkább a belső célok motiválják a szereplőket, és még a használat mérésekor is a hagyományos indikátorokat (pl. látogatók száma) helyezték előtérbe és nem az olyan, minőségi információkat is szolgáltató módszereket, mint a felhasználók bevonását igénylő használhatósági és hozzáférhetőségi vizsgálatok.

A számszerűsíthető paraméterek közül a digitalizálási költségek megállapítása különböző nehézségekhez vezetett: a válaszadók nem kaptak segítséget ahhoz, hogy hogyan számolják ki a fajlagos költségeket és sok helyen nem is álltak rendelkezésre a számításokhoz szükséges adatok. A teljesítmény mérése hasonlóan gondokat okozott a legtöbb felmérés során, mert nem sikerült egyértelmű és jól használható mértékegységeket definiálni. A digitális kimenet mérőszámai (pl. fájlok, képek, oldalak, gigabájtok mennyisége) nagyobb összefüggésekben nem értelmezhető adatok, szerencsésebb inkább az eredeti dokumentumok egységeiben mérni a digitális verziókat is.

A digitalizálás nem pusztán technikai folyamat, hanem egy sor stratégiai, gyűjteménykezelési, pénzügyi és egyéb jellegű döntés és cselekedet, annak érdekében, hogy a kulturális örökséget őrző intézmények jobban tudják teljesíteni küldetésüket. Ennek a komplexitásnak a megértése nem tükröződik a digitalizálás ellenőrzésére irányuló eddigi törekvésekben, aminek az az eredménye, hogy a különböző részterületekről csak töredékes információk gyűlnek össze, az egyes felmérések adatai pedig nem összevethetőek. A digitalizálási projektek hatékonyabb felügyeletéhez nagy szükség volna olyan szabványokra, amelyek mindenfajta gyűjteménytípusra alkalmazhatók és összehasonlítható eredményekre vezetnek nemcsak egyes intézmények, hanem a könyvtári, levéltári és múzeumi ágazatok között is, országos és nemzetközi szinten egyaránt. Ennek érdekében az Európai Bizott-

ság 2007-ben egy kétéves projektet indított *NUMERIC* néven. A projekt célja egy olyan monitoring rendszer kereteinek megalkotása, amely az illetékes nemzeti minisztériumok és más felügyeleti szervek képviselőinek bevonásával lehetővé teszi a nemzetközi statisztikai ügynökségeknek (pl. *EUROSTAT*, *UNESCO Institute for Statistics*), illetve egyes kulturális kutatóintézeteknek (pl. *European Institute for Comparative Cultural Research*), hogy rendszeresen begyűjthessék a digitalizálással kapcsolatos adatokat.

/MANŽUCH, Zinaida: Monitoring digitisation: lessons from previous experiences. = Journal of Documentation, 65. köt. 5. sz. 2009. p. 768–796./

(Drótos László)