

Digitális-hasonmás elismerő pecsét: egy felhasználó-orientált szabvány

Miért van rá szükség?

Az analóg hordozókon megjelent dokumentumok digitalizálása számos előnnyel jár mind a könyvtárak, mind az általuk szolgált közönség számára: az elektronikus formátum megkönnyíti a hozzáférést, a megőrzést, a megtalálást, a felhasználást, a gyűjteményszervezést... A digitalizálás azonban nem egy jól behatárolt, egységes folyamat, amelynek eredményeként automatikusan előállnak a felsorolt előnyök. Minden ilyen projekt egyedi: a rendelkezésre álló technológia, a költségvetés, a szakembergárda, a választott műszaki megoldások, és – jó esetben – a megcélzott felhasználói kör igényei alakítják ki azokat a körülményeket, amelyek között az adott digitalizálási munka megvalósul. A digitalizálás többféle dolgot is eredményezhet: fényképezés vagy szkennelés útján előállított oldalképeket, géppel felismertetett OCR-es vagy pedig emberi munkával begépelte szövegeket, de akár egy teljesen újragyártott e-könyvet vagy számítógépes alkalmazást is. És az analóg objektum digitálissá alakítása még csak az első lépés; más tevékenységekre (pl. metaadatolásra, megbízható repozitóriumban való megőrzésre, a digitális tartalomra épített szolgáltatásokra) is szükség van ahhoz, hogy a digitalizálás előnyei érvényesüljenek.

A „digitalizálás” szó jelentése tehát homályos, többet takar el, mint amennyit elárul. Hiába tudjuk, hogy egy könyvet digitalizáltak, ez még nem elég információ a könyvtárosok és a felhasználók számára, mert nem derül ki belőle, hogy milyen célra alkalmas az elektronikus verzió. Fontos volna, hogy ne csak a digitalizációs projektek irányítói, hanem a könyvtárak vezetői, a gyűjteményszervezéssel és a megőréssel foglalkozó szakemberek, a szolgáltatók és más érintettek – beleértve a könyvtárak használóit is – tisztában legyenek azaz, hogy egy adott digitalizálásnak mi lesz a tényleges eredménye, mi valósul meg igazán a lehetséges előnyök közül. Ez nemcsak a saját állomány digitalizálásakor fontos, hiszen a könyvtárak egyre inkább felhasználják és beépítik a gyűjteményükbe a máshol digitalizált, nyilvánosan elérhető vagy kereskedelmi forgalomba került anyagokat is.

Sokféle digitalizálási szabvány és útmutató létezik (a képi dokumentumokhoz pl. a FADGI Still Image Working Group állított össze róluk egy listát [1]), de

ezek elsősorban a digitális változat előállítására vonatkoznak és nehezen érthetők, nem igazán hasznosak a felhasználóknak vagy azoknak a könyvtárosoknak, akik a „végterméket”, vagyis a digitális példányt használni, illetve szolgáltatni fogják. Az olyan technikai jellemzők mint a színmélység vagy a felbontás, nem mondanak semmit a digitalizálás teljességéről és pontosságáról, mint ahogy arról sem, hogy egy adott felhasználási formára alkalmas-e a számítógépes változat. A jelenleg létező digitalizálási szabványokban nincs olyan „szótár”, amelyben leírható lenne az, hogy mik voltak az adott projekt céljai, hogy miért az adott megoldásokat választották, hogy sikerült-e elérni a célokat és a létrejött digitális objektumok megfelelnek-e a felhasználók igényeinek. Az előállításra vonatkozó szabványok kiegészítéseként tehát szükség van egy olyan világos, egyértelmű és objektív „fogyasztói” szótárra, amellyel a nemszakemberek is leírhatják és felhasználói szempontból összehasonlíthatják a digitalizálások eredményeit.

A bevezetésre javasolt *Digital-Surrogate Seal of Approval* (DSSOA) minősítés egy lépés ebbe az irányba. A DSSOA mint egyfajta elismerő vagy hitelesítő „pecsét”, azt jelzi, hogy a digitális változat pontos és teljes értékű helyettesítője az eredeti statikus, analóg objektum tartalmának és külalakjának.

Az olyan nyomtatott információcsomagok, mint amilyenek például a könyvek, az újságok és a folyóiratok, nem csupán tárolják és közvetítik az információkat, hanem kódolják és prezentálják is őket. Egy papírkiadvány egyben egy felhasználói felület is: a szöveg, a képek és egyéb elemek külalakja és elrendezése az oldalakon jelentést hordoz, kontextust ad. A tipográfia egyszerre tudomány és művészet, hosszú évszázadok alatt kialakult szokásokat követ. Egy külalakjától megfosztott szöveg akár az értelmét is elveszítheti, vagy legalábbis többértelművé, zavarossá válhat. A digitalizálás egyik fontos szempontja és célja ennek a kontextusnak, tehát az analóg példány külalakjának megőrzése és egy digitális szurrogátum – helyettesítő hasonmás – előállítása. Az ilyen célú projektek eredményeinek minősítésére szolgál a DSSOA szabvány. (Természetesen vannak olyan esetek is, amikor az eredeti külalak megőrzése nem cél, sőt kifejezetten hátrány, például mobil

eszközökön kényelmesen olvasható e-könyv készítése. Ezekre más minősítőket lehet majd bevezetni.)

Hogyan használható?

A DSSOA két kritériumot fogalmaz meg, valamint néhány szabályt ezek alkalmazására. A kritériumok a felhasználók számára a technológiai háttér ismerete nélkül is könnyen érthető fogalmak: *teljesség* és *pontosság*. A szabályok pedig a digitális másolatok létrehozásával, kezelésével, megőrzésével vagy terjesztésével foglalkozóknak adnak útmutatást annak megállapításához, hogy egy digitális dokumentum vagy egy dokumentumgyűjtemény megérdemli-e a DSSOA elismerő pecsétjét. Ha egy könyvtár úgy dönt, hogy állományvédelmi vagy gyűjteményszervezési okból az olvasóinak inkább a digitális példányt szolgáltatja, akkor a DSSOA megfelelő eszköz lehet annak jelzésére, hogy a digitális változat valóban alkalmas-e az eredeti helyettesítésére.

A DSSOA az ALA (*American Library Association*) digitalizálási ajánlásán [2] alapszik. A digitális hasonmástól elvárt két alapkövetelményt így határozza meg:

Teljesség: Az eredeti dokumentum minden oldala teljesen és tökéletesen reprodukálva van. Nincs olyan oldal vagy oldalrészlet, amely homályos, vagy elmosódott, vagy eltakart, vagy kihagyott, vagy hiányzik. Minden egyes oldal teljes egészében digitalizálva lett.

Pontosság: Az eredeti elrendezés és külalak meg lett őrizve. Minden szöveges rész olvasható, és legalább annyira könnyen, mint az eredeti példányon. Nincs látható minőségromlás. Az eredetihez hasonló méretben minden szöveg tiszta, elmosódás vagy torzulás nélküli. Valamennyi képi elem tiszta és a minősége (pl. szín, felbontás) az eredetivel megegyező.

Az alkalmazásra vonatkozó szabályok pedig a következők:

A DSSOA jelzés akkor használható egy bibliográfiailag azonosítható digitalizált tételnél, ha

- egy illetékes szervezet ellenőrizte, hogy mind a két kritérium teljesül az adott dokumentumnál;
- ha ez az illetékes szervezet kiadott erről egy *Megfelelőségi Igazolást (Statement of Verification)*.

A Megfelelőségi Igazolásban:

- meg kell adni és le kell írni a kritériumok ellenőrzéséhez használt módszert;
- ki kell jelenteni a 100%-os megfelelést.

Ha egy bibliográfiailag azonosítható tételből álló digitális gyűjtemény minden egyes darabja megfelel a kritériumoknak, akkor a DSSOA minősítés a gyűjteményre is alkalmazható.

A DSSOA szabvány nem határoz meg semmilyen konkrét módszert vagy eljárást az ellenőrzésre vonatkozóan. Ez történhet emberi közreműködéssel, vagyis az oldalak egyenkénti szemrevételezésével, de elfogadhatók automatizált megoldások is. Az igazolásban vagy le kell írni az alkalmazott eljárást, vagy hivatkozni kell egy olyan dokumentumra, amely ismerteti a teljesség és a pontosság ellenőrzésére használt módszert. Az ellenőrzésnél a reprezentatív mintavétel nem elegendő, minden egyes oldalnál – egy gyűjteményben pedig minden egyes tételnél – külön meg kell vizsgálni a kritériumok teljesülését.

A DSSOA-megfelelés jelzésére szintén nincs előírás a szabványban; a könyvtárak maguk alakíthatják ki, hogy hogyan tüntetik fel ezt az információt a metaadatok között. Egy ajánlott megoldás:

DSSOA_complete [yes | no ; date = ; verifier = ; notes =]

DSSOA_accurate [yes | no ; date = ; verifier = ; notes =]

DigitalSurrogateSealOfApproval [yes | no ; verification_method = ; qualification = [yes | no] ; notes =]

Bár a teljességi kritérium megköveteli, hogy minden oldal reprodukálva legyen a digitális változatban, amelynek valamilyen funkciója van (vagyis az üres oldalak is, ha pl. az oldalszámolás vagy a bal/jobbs oldalpárok megőrzése miatt szükségese), indokolt esetben – ha sem a kontextus miatt, sem jelentéshordozóként nincsen rájuk szükség – kihagyhatók oldalak a digitalizálásból, de akkor ezt a hiányt a *Megfelelőségi Igazolásban* jelezni kell. Az viszont nem elfogadható, ha azért maradnak ki egyes lapok, mert a projekt nem rendelkezik a digitalizálásukhoz szükséges eszközzel (pl. nagyméretű szkennelvel a kihajtható oldalakhoz vagy térképekhez). Az ilyen dokumentumok nem kaphatnak DSSOA pecsétet. Ha az eredeti analógkiadvány hiányos vagy megrongálódott és nincs

helyette hibátlan példány a digitalizálandó gyűjteményben, akkor a digitális változat megkaphatja a DSSOA jelzést, de a hiányt/hibát fel kell tüntetni az igazolásban, a metaadatoknál pedig egy minősítővel lehet ezt jelezni, például így: *Qualified Digital-Surrogate Seal of Approval: some content [may be / is] damaged or missing*. Hasonlóképpen kell eljárni, amikor a dokumentum egy másodlagos hordozóról (pl. mikrofilmről) lett digitalizálva, amely már eleve nem teljes és tökéletes másolata az eredeti papírkiadványnak. Ugyancsak minősítővel jelezhető gyűjteményszinten az, ha a digitális gyűjtemény egyes darabjai csak a minősített DSSOA-nak felelnek meg, például így: *Qualified Digital-Surrogate Seal of Approval: some copies are damaged or incomplete*. Ha utólag derül ki egy digitális dokumentumról, hogy mégsem teljesíti a követelményeket, akkor vagy vissza kell vonni a DSSOA jelzést, vagy korigálni kell a hibát és újra elvégezni az ellenőrzést. A gyűjtemény esetében pedig vagy ki kell venni a feltételeknek nem megfelelő tételt a többi közül, vagy a gyűjtemény tovább nem viselheti a DSSOA minősítést.

Mire használható?

A DSSOA kritériumokat arra használhatják a digitális tartalmakat létrehozó, kezelő, őrző vagy szolgáltató szervezetek, hogy tudassák felhasználóikat azok teljességéről és pontosságáról. Alkalmasak arra is, hogy egy digitalizálási projekt tervezési fázisában ezek figyelembevételével fogalmazzák meg az elérni kívánt célt: meg akarnak-e felelni ezeknek az elvárásoknak? A már létező digitális dokumentumok vagy gyűjtemények értékelői pedig javasolhatják, vagy megerősíthetik, vagy éppen megkérdőjelezhetik a DSSOA megfelelést. A felhasználók számára is hasznosak lehetnek ezek a fogalmak a digitalizált dokumentumok jellemzésére, vagy például különböző digitális verziók összehasonlítására abból a szempontból, hogy azok mennyire alkalmasak az eredeti analóg példány helyettesítésére.

A DSSOA nem használható mindenféle digitális projekthez vagy gyűjteményhez. Elsősorban olyan esetekben érdemes alkalmazni, amikor a digitális változatot a nyomtatott kiadvány helyett kívánják használni, nem pedig annak kiegészítőjeként. Emellett azoknak a könyvtáraknak is hasznos lehet, amelyek kereskedelmi és nonprofit forrásokból szereznek be digitális tartalmakat és szolgáltatják őket az olvasóiknak. Ilyen környezetben a DSSOA jó mérce az alternatívák közötti választáskor, valamint hasznos eszköz a választási lehetőség közléséhez.

A digitalizálás, a gyűjteményszervezés, a megőrzés és a szolgáltatás összetett folyamatok, egyetlen szabvánnyal nem lehet ezek minden aspektusát (pl. metaadatok, hosszú távú megőrzés, használhatóság, válogatási szempontok) egységesen leírni. A könyvtáraknak sokféle eszközre van/lenne szükségük, melyeket felhasználhatnak a döntéshozásnál. Ezek egyike lehet a digitális hasonmások minősítésére alkalmas DSSOA.

Hivatkozások

- [1] FADGI. Still Image Working Group: A Resource List for Standards Related to Digital Imaging of Print, Graphic, and Pictorial Materials. Federal Agencies Digitization Guidelines Initiative. 2010. január 28. <http://www.digitizationguidelines.gov/guidelines/digitize-standards.html>
- [2] BOGUS, Ian et al.: Minimum Digitization Capture Recommendations (DRAFT). Association for Library Collections and Technical Services Preservation and Reformatting Section. 2012. <http://www.ala.org/alcts/resources/preserv/minimum-digitization-capture-recommendations>

/JACOBS, James A. – JACOBS, James R.: *The Digital-Surrogate Seal of Approval: a Consumer-oriented Standard*. = *D-Lib Magazine*, 19. évf. 3–4. sz. 2013./

(Drótos László)

Robbanásszerűen fejlődik a beszéd felismerés

Pedro Moreno vezeti a Google Android részlegénél a beszéd irányítási és -felismerési fejlesztéseket. A szakember két évtizede foglalkozik a két szakterülettel.

„Nagyon érdekes fejleménynek tartom, hogy a beszéd felismerés hirtelen a középpontba került. Én 20 éve kutatom ezeket a rendszereket. Ez idő alatt volt egy tudományos párbeszéd és sor került számos kutatásra, de három évvel ezelőttig alig érdekelt valaki a terület iránt. Olyan téma volt, amiről senki sem gondolta igazán, hogy széles kör